Clinical, genetic and functional characterization of intellectual disability disorders

Christiane Zweier

The studies presented in this thesis were performed at the Department of Human Genetics, Radboud University Nijmegen Medical Centre, The Netherlands, and at the Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany.

The research presented in this thesis was financially supported by grants from the Deutsche Forschungsmeinschaft (DFG), from the German Federal Ministry of Education and Research (BMBF) as a part of the National Genome Research Network (NGFNplus), and from the Netherlands Organisation for Scientific Research (NWO).

Cover: Michael Kraft and Christiane Zweier

Printed and bound by: Gildeprint Drukkerijen, Enschede

ISBN: 978-3-00-046154-5

Clinical, genetic and functional characterization of intellectual disability disorders

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 22 september 2014
om 10.30 uur precies

door

Christiane Gertrud Zweier
geboren op 21 juni 1978
te Bamberg (Duitsland)

Promotoren:

Prof. dr. H.G. Brunner

Prof. dr. A. Reis (Friedrich-Alexander-University Erlangen-Nürnberg, DE)

Copromotor:

Dr. A. Schenck

Manuscriptcommissie:

Prof. dr. ir. J.H.L.M. van Bokhoven (voorzitter)

Prof. dr. M.A.A.P. Willemsen

Prof. dr. Y. Elgersma (Erasmus MC, Rotterdam)

Clinical, genetic and functional characterization of intellectual disability disorders

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. S.C.J.J. Kortmann,
according to the decision of the Council of Deans
to be defended in public on Monday, September 22, 2014
at 10.30 hours

by

Christiane Gertrud Zweier
born on June 21, 1978
in Bamberg (Germany)

Supervisors:

Prof. dr. H.G. Brunner

Prof. dr. A. Reis (Friedrich-Alexander-University Erlangen-Nürnberg, DE)

Co-supervisor:

Dr. A. Schenck

Doctoral Thesis Committee:

Prof. dr. ir. J.H.L.M. van Bokhoven (chairman)

Prof. dr. M.A.A.P. Willemsen

Prof. dr. Y. Elgersma (Erasmus MC, Rotterdam)

Contents

Chapter 1

General Introduction

1.1. Intellectual disability

Intellectual disability (ID), formerly mental retardation, affects 2%-3% of the population in western countries and is characterized by significant limitations in both intellectual functioning and adaptive behavior that begin before the age of 18 years and are usually reflected in an IQ below 70.[1,2] Severe or profound ID, reflected by an IQ below 35 or in simpler classifications below 50, affects 0.3% to 0.5% of the population.[3] The causes for severe ID are highly heterogeneous. Despite an increasingly known number of underlying genes, the genetic etiology still remains unsolved in nearly half of the cases.[4] Mild forms of ID are assumed to represent the lower end of normal IQ distribution and to result from the interaction of various genetic and other factors.[2] ID disorders represent a large socioeconomic burden for families and an economic burden for health care.[2]

1.2. Genetic causes of intellectual disability

The first genetic anomalies to be detected as causative for ID were numeric chromosomal aberrations. Trisomy 21 as the first one was identified in 1959.[5,6] In the 1970s banding techniques allowed the detection of cytogenetically visible structural chromosomal aberrations up to a resolution of 5-10 Mb in a diagnostic setting. With the development of techniques such as FISH (fluorescence in situ hybridization) in 1988,[7] MLPA (Multiplex Ligation-dependent Probe Amplification) in 2002,[8] and molecular karyotyping around 2000,[9-11] submicroscopic aberrations below 5 Mb and up to a resolution of several kb became stepwise detectable. Additionally, with the identification of single ID associated genes in the 1990s,[12] screening for point mutations with Sanger sequencing become more and more important. However, this allowed only analyses of single or a limited number of genes. Now, next generation sequencing (NGS) techniques are being established to identify point mutations by screening nearly the complete exome.[13] Progress in technology has thus stepwise increased our knowledge on the (different types of) genetic causes of diseases in general, and of ID specifically.

1.2.1. Chromosomal aberrations

Cytogenetically visible chromosomal aberrations can be found in ca. 16% of ID patients, with trisomy 21 still being the most frequent one.[14] With conventional karyotyping only numeric aberrations or structural aberrations larger than 5-10 Mb are detectable.

For many years FISH-analysis and more recently MLPA-analysis were used to detect smaller chromosomal aberrations. With these, common microdeletions or subtelomeric rearrangements could be identified in ca. 5% and 1.3% of patients, respectively.[14] However, FISH- and MLPA-probes are limited to specific chromosomal regions like subtelomeric

regions or previously known microdeletion regions. The development of molecular karyotyping by arrayCGH or SNP arrays allowed genome-wide screening and a resolution of currently 100 kb in a diagnostic setting.[15] It therefore increased the diagnostic yield of chromosomal aberrations in patients with ID and previously normal karyotype to ca. 10-15%. The detection rates range between 5.5 - 35%, depending on the resolution of the used array and the pre-selection of patients.[16]

Chromosomal aberrations can occur as deletions or duplications with variable breakpoints and sizes. This is observed in many rare, individual aberrations or in the more common 4p- or Wolf-Hirschhorn syndrome[17] and the 1p36 microdeletion syndrome.[18] Furthermore, they can occur as common microdeletions with identical and recurrent sizes and breakpoints, usually mediated by low copy repeats resulting in non-allelic homologous recombination.[19,20] This applies for example to DiGeorge syndrome, caused by a typical deletion in 22q11.2[21] or to Williams-Beuren syndrome, caused by a recurrent 1.55 (95%) or 1.84 Mb (5%) deletion in 7q11.23.[22]

Molecular karyotyping not only increased the findings of individual small chromosomal aberrations but also resulted in the identification of novel, relatively common microaberration syndromes like the 17q21.31 microdeletion syndrome[23-25] or the 16p11.2 microdeletion/duplication syndrome,[26] thus adding to the number of previously known common microaberration syndromes such as Di George syndrome, Williams-Beuren syndrome, and Smith-Magenis syndrome.

Microdeletions or –duplications can act as contiguous gene syndromes with several deleted/duplicated genes contributing to the phenotype. This applies e.g. for Williams-Beuren syndrome or monosomy 1p36.[18,27] For other microdeletion syndromes, responsibility of a single deleted, dosage sensitive gene, was identified. Regarding the phenotype of monosomy 22q11.2, *TBX1* is considered as the most promising candidate gene,[28] and for the 9q23 subtelomeric microdeletion, *EHMT1* was identified as the phenocritical gene.[29] In both genes point mutations were shown to result in a similar phenotype as the deletions.[29-31]

The vast majority of larger chromosomal aberrations and copy number variants (CNVs) occurs *de novo* as reproduction of (severely) affected patients is usually limited. However, microaberrations that are associated with mild forms of ID are often found segregating in families, e.g. monosomy 22q11.2[32] and deletions or duplications of 16p11.2.[33]

1.2.2. Monogenic causes

To date, about 518 genes are reliably implicated in ID. Ca. 59% harbor autosomal recessive and ca. 24% harbor autosomal dominant mutations. Further 20% are located on the X-chromosome and 1.7% in the mitochondrial genome (status May 2013, see chapter 6 of this

thesis). Two years ago the total number was still estimated to run into the thousands.[2] However, as recent studies in patients with unspecific severe ID showed a mutation detection rate of 13-30% in previously known ID genes,[34-36] the final number might be not as large as proposed.

Fragile X syndrome is considered to be the first identified ID syndrome caused by a genetic defect in a single gene.[37] It is also the most frequent cause for monogenic ID.[38] With more than 90, a relatively large number of X-linked ID genes is known to date. Gene identification for X-linked recessive ID has been facilitated by the inheritance pattern. This is often characterized by several affected males and unaffected carrier women, thus allowing linkage analysis.[39] Due to the limited number of genes on the X-chromosome, also large-scale systematic re-sequencing of all coding exons could be performed.[40] In total, genetic defects on the X-chromosome are supposed to underlie ID in 8-12% of affected males.[41] In addition, as for example regarding Rett syndrome, some X-linked dominant mutations are known. These are usually *de novo*, accompanied by a phenotype in females and assumed to be lethal in males.

This leaves a large number of supposedly autosomal genetic defects. Gene identification for autosomal recessive disorders was until recently easier than for mostly sporadic, dominant disorders. Many recessive metabolic disorders show specific phenotypes, and gene identifying methods such as linkage analysis only work in inherited disorders. How many genetic defects underlying ID will in the end be autosomal recessively inherited or will be sporadic dominant *de novo* mutations, is therefore difficult to estimate. However, recent studies indicate that recessive mutations in outbred populations seem to be rather rare in contrast to *de novo* mutations.[34,35,42]

In disorders with severe ID and therefore limited reproduction, the causative autosomal dominant or also X-linked dominant mutations are usually found to have occurred *de novo*. In contrast, dominant mutations that are associated with rather mild or low penetrant cognitive defects can be observed to be inherited in families. This applies for example to Noonan syndrome.[43]

Mutations in mitochondrially encoded genes can also be associated with ID in the context of mitochondrial disorders, however, compared to X-linked and autosomal dominant causes of ID they are rare. Mutations in the mitochondrial genome either occur *de novo* or are transmitted by the mother.[44]

1.2.3. Others

In mild cases of ID or in autism spectrum disorders (ASD), which are usually accompanied by rather mild cognitive impairment, the underlying genetic defect might be more complex.

This means that a single mutation or CNV represents a risk factor which is only pathogenic in combination with one or more additional "hits". This has been shown for the 16p12.1 microdeletion,[45,46] but is also discussed for defects in single genes such as *SHANK3*[46] or *FOXP1*.[47] For some disorders such as Bardet-Biedl syndrome, a di- or even trigenic inheritance has been reported.[48] In general, the understanding which combination of multiple hits underlie certain neuropsychiatric disorders and how they mutually interact is currently still at the beginning and significantly more difficult than for monogenic disorders.

Some forms of mutation do not affect genes themselves, but their regulation. This is the case in imprinting disorders. Deregulation of imprinted genes can for example result in Angelman- or Prader-Willi syndromes, depending on, whether the maternal or paternal copy is lacking or inactive.[49]

Very little has been known about mutations or other aberrations in regulatory elements of the genome so far. While CNVs in non-transcribed regions are reported in several limb malformation syndromes (reviewed by Klopocki and Mundlos[50]), their relevance for ID disorders is only known for single cases so far. For example, an intronic deletion in the PLP gene can lead to an abnormal presentation of Pelizaeus-Merzbacher disease,[51] and a non-coding mutation regulates *HCFC1* expression in non-syndromic ID.[52] Aberrations in non-coding elements and their pathomechanisms represent a level of complexity whose extent is currently difficult to foresee.

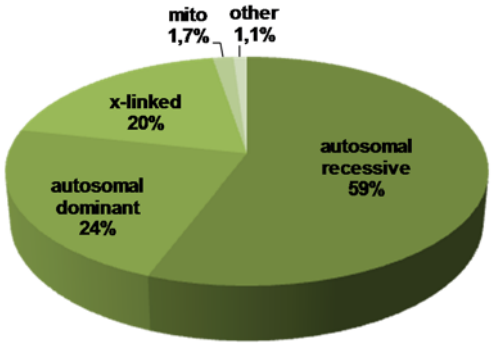

Figure 1 Distribution of the inheritance pattern of 518 genes reliably implicated in ID.
(status May 2013, see chapter 6 of this thesis). Percentages refer to the proportion of genes. Mito, mitochondrially encoded.

1.3. Identification of novel ID genes

Dependent on the occurrence of ID in several members of a family or sporadically due to a *de novo* event, different strategies for gene identification can be used. While methods such as linkage analysis or breakpoint mapping have in principle been available for more than two decades, they developed their real power only during the last decade. SNP arrays and the

evolution of sequencing technologies allowed high resolution homozygosity mapping and an increased screening throughput of candidate genes, respectively. As to familial ID, a combination of mapping methods with targeted re-sequencing of candidate regions by recently developed NGS technologies can be considered as the optimization of previously available strategies. Regarding sporadic ID, NGS can be considered as the big, revolutionary step forward. For the first time a systematic approach to unravel its underlying genetics can be undertaken.

1.3.1. Linkage analysis

Linkage analysis can be used when an ID disorder segregates within a family and when several affected and unaffected individuals are available for testing. Through positional cloning, genes following mendelian traits can be identified based only on the knowledge that the phenotype is inherited. Further information on the biology of the disease or its cause is not necessary.[53] Linkage analysis uses Linkage Disequilibrium (LD) within families with a segregating phenotype to identify co-segregating haplotype blocks. To achieve this, DNA polymorphisms like previously microsatellite markers and now single nucleotide polymorphisms (SNPs) contained on microarrays[53,54] are utilized. To obtain sufficient LOD scores, reflecting the likelihood of linkage between two genetic traits, large families or several smaller families with the same phenotype and overlapping candidate regions are required. Until recently, this was the bottleneck in disease gene identification as time and cost consuming Sanger sequencing allowed screening only of a limited number of candidate genes. With (targeted) NGS larger regions containing a higher number of genes can be screened.

Linkage analysis has been successfully applied to autosomal dominant ID associated disorders such as Neurofibromatosis type I,[55,56] to autosomal recessive ID syndromes like Cohen syndrome,[57] and to X-linked recessive ID caused for example by mutations in the oligophrenin gene.[39]

1.3.2. Homozygosity mapping

Homozygosity mapping works only for recessive disorders in consanguineous families and is based on the assumption that the phenotype is caused by a true homozygous mutation. A single heterozygous variant for a recessive disease has a very high chance of becoming homozygous and therefore disease causing in the offspring of consanguineous parents.[58] Similar to linkage analysis, mutational screening in candidate genes within homozygous regions is performed. Systematic studies using homozygosity mapping require populations with a high degree of parental consanguinity and large family sizes accompanied by a high

frequency of ID and congenital anomalies. Such conditions are typically found in Arab countries, Turkey, Iran, Pakistan, and some parts of India.[59,60] Large studies in families from Iran or Syria using DNA array-based SNP typing revealed various novel loci for non-syndromic autosomal recessive ID.[61,62] In the biggest study with 136 families this approach led to the identification of mutations in 23 genes previously implicated in ID or related neurological disorders. Probably disease-causing variants were additionally identified in 50 novel candidate genes.[63] These studies therefore emphasize the extreme heterogeneity of autosomal recessive ID.

1.3.3. Chromosomal translocations

De novo balanced chromosomal translocations can be identified by conventional karyotyping and occur with an incidence of 1 in 2000.[64] Six percent of them are associated with ID with or without further anomalies.[60] Exact breakpoint mapping of such translocations can therefore elucidate candidate genes for ID.[65] Subsequent confirmation can be achieved by identifying mutations in the same gene in other patients with a similar phenotype. This approach has been successful for autosomal genes like *ZEB2,* mutations in which underlie Mowat-Wilson syndrome,[66] or for *GRIN2A* and *GRIN2B,* mutations in which cause variable ID with or without epilepsy.[67] Furthermore, this approach was also successful for the X-linked dominant *CDKL5* gene.[68,69]

Breakpoint mapping cannot be pursued systematically but depends on the finding of apparently balanced chromosomal translocations.

1.3.4. Molecular karyotyping by arrayCGH and SNP arrays

Apart from an increasing diagnostic outcome in patients with ID harboring individual small chromosomal aberrations[70,71] or by identifying novel common microdeletions syndromes (reviewed by Slavotinek[72]), molecular karyotyping by arrayCGH or SNP arrays has been proven to be a suitable method to identify novel disease causing genes. After initial identification of small aberrations affecting only one or few genes, subsequently mutations in the respective candidate gene in further patients without copy number variants can be uncovered. This applied to well-known syndromic disorders such as autosomal-dominant CHARGE syndrome[36] or autosomal recessive Peters Plus syndrome.[73] Further examples are the identification of causative genes for an epileptic encephalopathy[74] or for non-specific moderate to severe ID.[75,76]

1.3.5. Sequencing of functional candidate genes

Targeted screening for mutations in a particular gene by conventional Sanger sequencing is often the method of choice if this gene is considered to be a good ID candidate gene. This can for example be due to its known function or due to an interaction with known disease causing genes/proteins.

Selecting a functional convincing candidate gene can be done in combination with linkage analysis and homozygosity mapping (see above), where it can be picked from a larger but still limited number of genes. Furthermore, several studies proved the feasibility of selecting and screening a candidate gene purely on its function without additional genetic information. Within the framework of the Canadian Synapse to Disease consortium, Hamdan et al. found *de novo* protein-truncating mutations in *SYNGAP1*[77] in patients with non-syndromic ID as well as in other synaptic genes associated with glutamatergic systems.[78] *SYNGAP1* had been known to be involved in pathways regulated by NMDA receptors, and mice with heterozygous Syngap1 mutations had been found to have cognitive dysfunction.[77]

Mutation screening for *RPGRIP1L* was performed because of its interaction with ciliary and basal body proteins involved in Joubert syndrome and indeed led to the identification of mutations in patients with a typical Joubert syndrome phenotype.[79]

1.3.6. Next generation sequencing

Sanger sequencing allows only screening of a limited number of genes, e.g. single genes or genes from linkage or homozygosity mapping regions or, in an exceptional large scale attempt, of all coding exons of the X-chromosome.[40] During the last few years NGS technology has evolved as the key to overcome these limitations.[80]

For the identification of the underlying mutations in monogenic disorders like many forms of ID, currently mainly whole exome sequencing is used. As only about 1% of the whole genome is protein coding, massive parallel sequencing of those regions has been become a relatively rapid and cost efficient method. Exome sequencing starts with enriching exonic sequences using different array-based or solution based capturing methods.[26] This is followed by massive parallel sequencing and a bioinformatics pipeline that filters the identified changes to distinguish pathogenic mutations from common variants or technical errors.[34]

Whole exome sequencing can be utilized for disease gene identification based on different hypotheses. In index patients of consanguineous families screening for recessive mutations in homozygous regions can be performed much more extensive than before with Sanger sequencing and resulted in the identification of several new ID related genes or candidate genes.[63,81,82] This also applies to families with X-linked ID and targeted NGS for

genes on the X-chromosome (Kalscheuer et al., Next-generation sequencing in 248 families with X-linked intellectual disability; Abstract #84, Presented at the 12th International Congress of Human Genetics/61st Annual Meeting of The American Society of Human Genetics, October 13th, 2011, Montreal, Canada).

The largest progress was made regarding sporadic *de novo* mutations. The first two ID syndromes solved by this method were Kabuki syndrome[83] and Schinzel-Giedion syndrome.[84] Several patients with a similar clinical phenotype were tested simultaneously and subsequently screened for mutations in the same gene. This approach relies on syndromic forms of ID where several patients with a recognizable similar phenotype are available. In non-syndromic ID the so called trio approach has been successfully used by testing patient and parents simultaneously and looking for *de novo* mutations in previously known or functionally convincing genes.[34,35,42]

Table 1 Strategies for disease gene identification

Strategy	Suitable for the indicated inheritance patterns	Specifics
linkage analysis	familial x-linked, autosomal recessive and dominant	currently NGS used for candidate gene screening in linkage regions
homozygosity mapping	autosomal recessive	mainly for consanguineous families
chromosomal translocations	sporadic, dominant; (familial dominant)	based on "incidental" findings, no systematic approach
molecular karyotyping	sporadic, dominant	CNVs of one or few genes of which candidate gene has to be selected
functional candidate genes	independent from inheritance pattern	based on function rather than on genetics
NGS	all	systematic, can be used unbiased

1.4. The clinical faces of ID

ID comes in many facets. It is defined as a significant limitation in both intellectual functioning and adaptive behavior starting before the age of 18 years (see 1.1.). This definition includes primary cognitive impairment that is often initially noted by delay of developmental milestones as well as disorders that are accompanied by no development at all, by early neurodegeneration and regression, and sometimes even by lethality.

Not only the cognitive and neurological performance in ID disorders can be highly variable, but also the accompanying phenotypes. Classically, ID is clinically distinguished into syndromic and non-syndromic forms. While non-syndromic ID disorders are not marked by any specific phenotypic aspect, characteristic features co-occurring with ID in syndromic disorders can be very variable.

Cornelia-de Lange syndrome and lissencephaly, for example, are characterized by a whole spectrum of distinct dysmorphism and malformations of extremities and organs in case

of the first and specific brain malformations in case of the latter.[85,86] Such structural malformations can be noted by external clinical examination of the patient, by ultrasound examination of the organs, and by MRI (magnetic resonance imaging) examination of the brain. Physical examinations are therefore an important first step in the diagnosis of syndromic ID as they can provide a clue on the disorder and prompt specific molecular testing.

Other disorders like Down syndrome, Williams-Beuren syndrome[87] or Mowat-Wilson syndrome[88] are associated with a specific facial gestalt that alone is sufficient to allow clinical diagnosis in most of the patients.

In some ID disorders the disease course over several months or years can be the main contributing factor for the clinical diagnosis. Typical Rett syndrome, for instance, is characterized by the sequence of normal development, stagnation, regression, and then long-term stability.[89]

Many of the metabolic ID syndromes are not characterized by any physical anomalies but by specific enzymatic, metabolic or biochemical alterations in blood, urine or liquor. Some of them can also lead to specific non-malformation anomalies detected with brain MRIs like for example in lipid storage disorders such as X-linked adrenoleukodystrophy.[90] Only few inherited metabolic diseases cause isolated stable ID. Additional neurological signs such as regression, ataxia, seizures, movement disorders or behavioral problems are more commonly found.[91]

Furthermore, there are disorders that can be accompanied by ID but without ID being a major aspect. This applies for example to Neurofibromatosis Type I, which is mainly characterized by skin features such as Café-au-lait macules and neurofibromas. But it is also associated with learning disabilities in 50 to 75% of patients.[92]

Some forms of ID were initially considered as non-syndromic and later re-classified to syndromic ID after the underlying cause was identified or after patients were analyzed in more detail. This happened for example with the AP4 syndrome. Specific features like shy personality and spasticity of the lower extremities were commonly recognized in several individuals carrying defects in the same gene or genes from the same complex.[82]

For the past years the prevalence of diagnosed autism spectrum disorders (ASD) has rapidly increased.[93] This is probably not only due to a real increase in co-morbidity but rather due to softening borders of terminology and by including ID patients into ASD studies, based on phenotypic overlap. The mixing of ID with neuropsychiatric disorders such as autism is reflected in the estimation that approximately 10% of children with ID have autistic symptoms and that 70% of individuals with autism also have ID.[94,95] This clinical overlap is probably at least partly based on a molecular overlap of ID and autism related molecular pathways,[96] but might also be influenced by changing diagnostic criteria and societal biases.[93]

1.5. The molecular faces of ID

Genetic defects can be pathogenic due a variety of molecular consequences. In principle, a genetic alteration can by several mechanisms result in loss of function or gain of function of the encoded protein, respectively.

1.5.1. Dosage sensitivity

The phenotype of trisomy 21 (Down syndrome) is assumed to result from an increased dosage of one or more of the ~310 genes present on this chromosome.[97] Several dosage sensitive genes have been identified that might play a role in the pathogenesis of Down syndrome, amongst them *DYRK1A*. Overexpression of *DYRK1A* results in severe learning deficits in mice,[97] and deletions are reported in association with ID and primary microcephaly.[98]

Also for *TBX1*, a gene contained in the 22q11.2 microdeletion and –duplication region, and in which point mutations were shown to cause the monosomy 22q11.2 phenotype, a time and tissue dependent dosage effect is discussed.[99] The *RAI1* gene is assumed to be the dosage sensitive gene in clinically distinct Smith-Magenis and Potocki-Lupski syndromes, caused by reciprocal microdeletions and –duplications in 17p11.2, respectively.[100-102]

1.5.2. Loss of function

Loss of function can be caused by different mechanisms that all result in inactivation of the affected gene/protein. Nonsense mutations, frameshift mutations, splice mutations and out-of-frame intragenic deletions or duplications can lead to premature termination codons. This can result in truncation of the protein with loss of important C-terminal domains as for example assumed for *KAT6B* mutations in Genitopatellar syndrome[103,104] and distinct Say-Barber-Biesecker Variant of Ohdo Syndrome.[105] It is suggested that loss-of-function, either by haploinsufficiency or loss of the c-terminal region, results in overlapping features between the two syndromes, whereas an additional dominant-negative or gain-of-function effect is responsible for the more specific symptoms.[104,106]

More commonly, aberrant mRNAs are subject to nonsense mediated mRNA decay,[107] leading to a reduction of mRNA and subsequently protein. Heterozygous loss-of-function mutations often result in haploinsufficiency (see 1.5.3.), while biallelic loss of function often affects enzymes, whose complete or nearly complete loss is often found in metabolic diseases.

Additionally, also missense mutations within important functional domains can cause loss of function as shown for example for MeCP2.[108]

1.5.3. Haploinsufficiency

Haploinsufficiency occurs, when a single functional allele is not sufficient to produce a normal phenotype. This mechanism typically underlies dominant heterozygous loss-of-function defects or the loss of dosage-sensitive genes, both described above.

Haploinsufficiency of a specific gene can be assumed, when larger deletions of a gene and point mutations within this particular gene are found to result in the same phenotype. This was, for instance, shown for Kleefstra syndrome that is either caused by a microdeletion of 9q34.3 containing the *EHMT1* gene or by mutations in the *EHMT1* gene itself.[29] Also larger deletions in 5q as well as point mutations – both stop mutations and missense mutations- in the *MEF2C* gene cause an identical phenotype of severe ID with epilepsy.[75]

1.5.4. Gain of function

Gain-of-function mutations result in increased or abnormal function of the gene product. Therefore gain of function can lead to more specific consequences compared to loss of function.

Noonan syndrome and related Rasopathies are caused by deregulation of the RAS-MAPK signaling pathway. For many mutations that were identified in positive regulators of the pathway (SHP2, SOS1, KRAS, HRAS, BRAF, MEK1, MEK2), a gain-of-function effect could be experimentally shown.[109] Most of these mutations are missense changes, and a few are deletions of single amino acids in particular functional domains of the respective protein.[109]

Also missense mutations in *SETBP1*, causing Schinzel-Giedion syndrome, and clustering in a highly conserved 11 bp region, are supposed to act in a gain-of-function manner.[84] Schinzel-Giedion syndrome is characterized by severe ID, distinctive facial features, multiple congenital malformations, and lethality within the first decade.[84] The specific effect of the supposed gain-of-function mutations in *SETBP1* in those patients is confirmed by the finding of *SETBP1* haploinsufficiency in patients with a different phenotype of mild ID and expressive language impairment.[110]

The heritability of intelligence: gain of function versus loss of function?

Though the heritability of intelligence is high, only very few genetic factors have been identified so far. Due to their number, the effect of each single factor is probably low.[111] It is suggested that genes associated with intelligence might also be of importance to ID. This means that variants in the same gene can contribute to both upper (normal or high

intelligence) and lower (ID) extremes of intelligence as a normally distributed trait.[112,113]
One might hypothesize a rather high proportion of loss-of-function defects in ID compared to gain-of-function variants which might contribute to normal or high intelligence. This idea is supported by a larger number of ID associated recurrent chromosomal deletions than duplications (15 versus 8, Vissers et al.,[114]). However, strategies in the pre-NGS era have been biased towards loss-of-function defects due to a high fraction of metabolic diseases and due to genes identified by chromosomal aberrations which mostly are deletions or gene-disrupting translocations.
Interestingly, for an increasing number of genes, gain-of-function as well as loss-of-function mutations have been implicated in ID. These can result in distinguishable phenotypes, respectively. This applies to *SETBP1* as described above[84] as well as to *SCN2A.* Missense mutations in this gene cause an early onset epileptic encephalopathy and were shown to have a gain-of–function effect.[115] In contrast, truncating mutations were recently identified in individuals with unspecific ID and autism without epilepsy.[34,35,116] It might therefore be suggested that gain-of-function mutations lead to a more complex and syndromic phenotype than loss-of-function mutations. The increasing findings of different phenotypes caused by mutations in the same gene might also implicate that at least ID associated with sporadic *de novo* mutations might be genetically less heterogeneous than proposed before.

1.5.5. Dominant negative effects

A dominant negative effect occurs when a mutant protein interferes with the function of the normal second allele in a heterozygous person.

Such an effect is discussed for mutations identified in the *SMARCA2* gene in patients with Nicolaides-Baraitser syndrome. All point mutations found so far have been missense mutations within the SNF2_N and HELICASE_C ATPase domains. Additionally, one deletion encompassing the C-terminal helicase domain was observed.[117,118] These findings, together with data from yeast, support a dominant-negative model in which dysfunctional but structurally undamaged SMARCA2 generates a complex that is intact with respect to its composition and occupies its appropriate chromatin domains but is nonetheless functionally inactive.[117]

1.5.6. Dynamic mutations

Fragile X syndrome, the most common cause of monogenic ID, is caused by an expanded CGG repeat in the promoter of the *FMR1* gene, leading to hypermethylation and therefore inactivation of the protein.[37]

Another example is the congenital form of myotonic dystrophy type I which is often accompanied by ID. This disorder is caused by a trinucleotide expansion in the non-coding region of the DM1 protein kinase gene, *DMPK,*[119] resulting in accumulating and thus possibly toxic RNA molecules.[120,121]

1.5.7. Imprinting defects

Genomic imprinting describes the preferential or exclusive expression of a gene on either the maternal or paternal allele. The allele-specific expression of imprinted genes is mediated by allele-specific epigenetic modifications such as DNA-cytosine methylation, histone acetylation, and other modifications.[49] Abnormalities in imprinted inheritance occur in several well-known developmental and neurobehavioral disorders, including Albright's hereditary osteodystrophy or Beckwith-Wiedemann syndrome.[122] Loss of a functional paternal or maternal allele in chromosomal region 15q11-q13 due to deletion, uniparental disomy, or mutations in imprinting centers, results in the clinically distinct phenotypes of Prader-Willi or Angelman syndromes, respectively.[122]

1.6. Common themes among ID disorders

Many of the genes and proteins implicated in ID can be assigned to different categories according to their function as enzymes, mediators of signal transduction, transcriptional regulators, etc. or according to their biological function in processes like nuclear organization, metabolic and signaling pathways, organization of cytosceleton.[4] Defects are suggested, apart from potential functions outside the nervous system, to lead to two major common groups of phenotypic outcomes: 1. dysfunctional neurodevelopment and brain malformation and 2. alterations in molecular mechanisms underlying synaptic organization and plasticity.[4] However, a complete summary as well as a comprehensive overview on all ID-related genes is still lacking. Several examples of genotypic and/or phenotypic entities are given here. However, they often show considerable overlap between each other and are by far not complete.

1.6.1. Inborn errors of metabolism

For many inborn errors of metabolism the underlying biochemical defect, often integrated in larger pathways, is well known. However, the exact mechanisms of brain damage and dysfunction are still poorly understood.[91,123]

Lysosomal storage disorders are characterized by deficient intracellular transport or metabolism resulting in progressive accumulation of un-degraded catabolites. This occurs

particularly in the reticuloendothelial system and the nervous system, thus often leading to progressive central nervous system damage by disturbance of myelin formation and function.[123,124]

Also phenylketonuria seems to result in white matter anomalies and dysmyelination by direct neurotoxic effects. These are discussed to result from elevated phenylalanine levels leading to disturbed amino acid transport into the brain and alterations in neurotransmitter levels.[123,125] Aberrant neurotransmission is also observed in glycine or serine disorders.[123]

Furthermore, deficiencies in energy availability as seen in creatine deficiency as well as defects in energy production caused by mitochondrial disorders can underlie cognitive impairment.[123] The respiratory (or electron transport) chain in mitochondria consists of 4 complexes comprising ca. 90 subunits which are mainly encoded by nuclear genes. It is the core machinery for oxidative phosphorylation and a hub in the cellular metabolism network. Disruptions can lead to a broad spectrum of phenotypes ranging from highly tissue-specific to multisystemic symptoms.[44]

1.6.2. Neurogenesis

The first and critical step of cortex development is the generation and proliferation of neuronal precursor cells. Various neurological conditions are associated with alterations in mechanisms of cell proliferation, cell fate determination or programmed cell death, thus leading to an abnormal number of neurons.[4] The resulting abnormalities in brain size are characteristic for primary microcephaly disorders. These are caused by mutations in genes like *MCPH1*,[126] *ASPM*,[127] or *CENPJ,*[128] which are involved in centrosome function and DNA repair response pathways.[4]

1.6.3. Neuronal migration

Neuronal migration disorders involve disruptions of various stages of the migration process, including the onset of migration, migration movements, the penetration into the preplate, and the arrest of migration.[4] Depending on the time, type and localization of migration defects, different brain malformation phenotypes can manifest. These comprise lissencephaly, heterotopias, polymicrogyria and schizencephaly.[129] Examples are mutations in *LIS1*, *DCX*, and *TUBA1A* in classical lissencephaly, mutations in *POMT1*, *POMT2*, *FKTN* etc. in cobblestone lissencephaly, mutations in *FLNA* or *ARFGEF2* in periventricular heterotopias, mutations in *GPR56* in polymicrogyria, and mutations in *EMX2* in schizencephaly (summarized and reviewed by Verrotti et al.[129]).

1.6.4. Synapse formation and plasticity

The formation, maintenance and modification of synapses are fundamental to learning, memory, and other cognitive processes. Mutations in the "synaptome" can result in a broad spectrum of neurodevelopmental and cognitive disorders.[130] However, despite being a major theme in ID the degree of its contribution is still debated.[63]

Synaptic plasticity is reflected by experience- or activity-induced changes in molecular composition and morphology of synapses and dendritic spines, the main sites of excitatory synaptic input. It is assumed as a correlate for learning and memory.[3] Various defects in presynaptic pathways, in postsynaptic protein complexes, or in the cytoskeleton can lead to impaired synaptic plasticity and therefore to cognitive dysfunction.[3,4]

ID genes with a role in presynaptic vesicle recycling are for example *STXBP1,*[74] *GDI1,*[131] and *RAB39B.*[132] Presynaptic CASK has an important role in trans-synaptic protein interaction and is implicated in severe ID with brain malformations.[133]

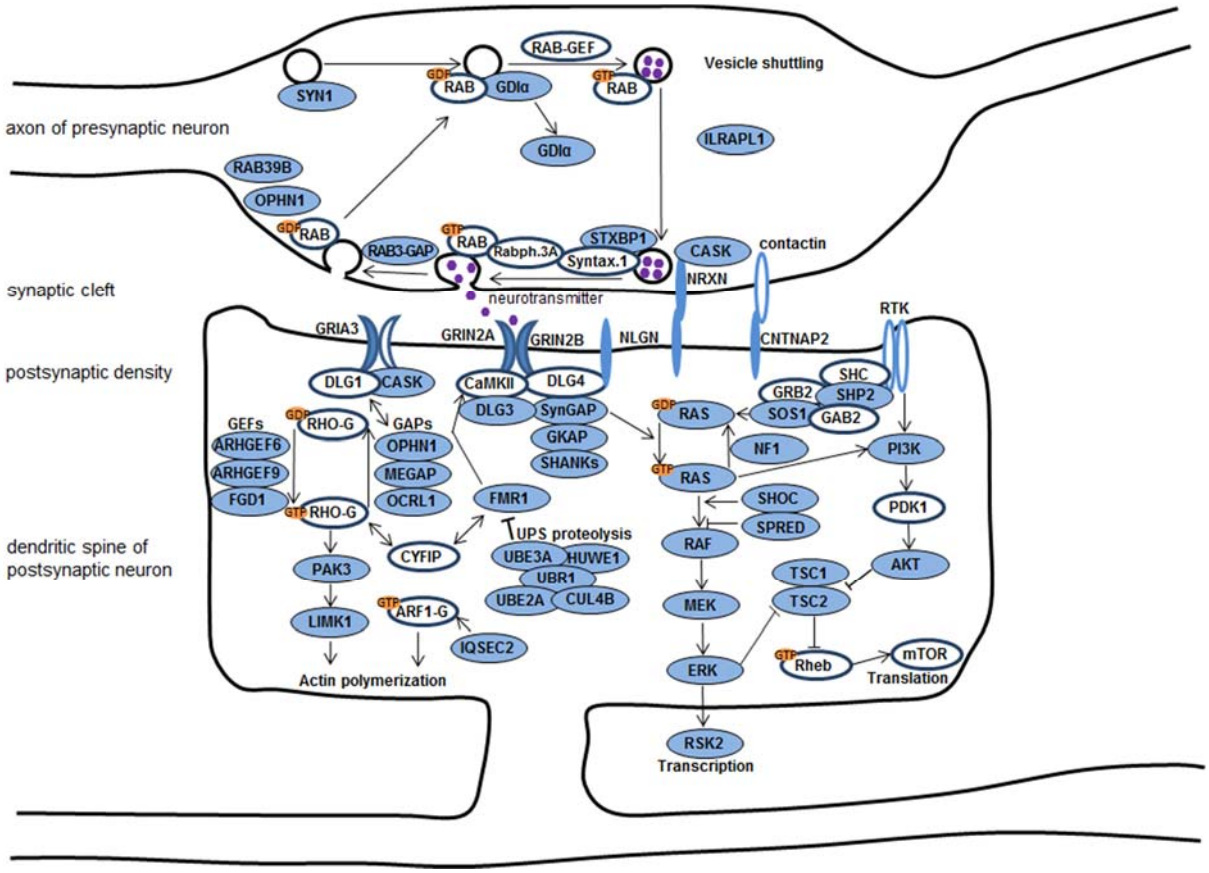

Figure 2 Schematic drawing of a synapse with some pre- and postsynaptic networks and pathways, involved in various processes.

Shaded proteins are encoded by known ID genes. Adapted from Van Bokhoven, 2011[3]. G., GTPase; Syntax.1, Syntaxin-1; Rabph.3A, Rabphilin 3A, RTK, receptor tyrosin kinase

ID associated genes involved in postsynaptic complexes are for instance *GRIA1*, encoding an AMPA receptor subunit,[134] or *GRIN2A* and *GRIN2B*, coding for NMDA receptor subunits.[67] Other proteins involved in postsynaptic density complexes are SYNGAP[77] and the scaffolding proteins SHANK1-3.[135-137] Also local regulation of postsynaptic protein levels seems to be critical for normal learning and memory processes.[3] This is demonstrated by mutations in the Fragile X syndrome gene *FMR1,* which encodes a regulator of local synaptic protein snythesis.[138] Another example are mutations in the Angelman syndrome gene *UBE3A,* which is involved in ubiquitin-proteasome system-dependent protein degradation.[139,140]

Members of RhoGTPase signaling pathways are important regulators of dendritic spine morphology and synaptic acitivity.[49] Deficits in the regulation of the dendritic cytoskeleton affect both the structure and function of dendrites and synapses, common features in patients with ID.[141] The RhoGTPase pathway includes ID genes such as *OPHN1,*[39] *PAK3,*[142] *ARHGEF6,*[143] and *FGD1.*[144]

Another well-known example for a cellular signaling cascade implicated in ID is the RAS-ERK signaling pathway with both presynaptic and postsynaptic roles.[3,145] It is involved in Noonan syndrome and related disorders and comprises the Ras GTPases HRAS and KRAS[146] as well as the downstream effectors RAF1, BRAF, MEK1, MEK2 and RSK2, and regulators of the Ras-MAPK pathway such as SHP2, SOS, NF1, SPRED1, and SHOC (reviewed by Tartaglia et al.[147]), thus highlighting a nearly complete molecular pathway. The associated disorders of the Noonan/CFC/Costello syndrome spectrum are often accompanied by cognitive deficits of variable severity, which seems to be at least partially correlating with the affected gene and the type of mutation.[148]

1.6.5. Transcriptional regulation

Transcription is, amongst other processes, a molecular prerequisite for long-term synaptic plasticity and long-term memory function.[149] It is controlled by transcription factors and co-factors, by signal transduction cascades, and by chromatin-remodeling proteins.[49]

Transcription factors bind to particular cognate sequences in promoter regions of their target genes and regulate their expression. Examples for transcription factors implicated in ID syndromes are ZEB2 in Mowat-Wilson syndrome,[150,151] MEF2C in severe Rett-like ID,[75,152] or SOX3 in X-linked ID with growth hormone deficiency[153] and SOX10 in the neurologic variant of Waardenburg-Shah syndrome.[154]

Furthermore, transcription is regulated by epigenetic modulation of chromatin structure. This includes mechanisms like DNA methylation, modification of histone proteins, and ATP-dependent chromatin remodeling.[155] The Rett syndrome gene *MECP2* encodes the methyl-CpG binding protein 2. It binds to methylated CpG dimer pairs in DNA and recruits

certain transcriptional co-factors like histone deacetylases. This subsequently leads to chromatin condensation and repression of target gene expression.[156,157] Mutations in *MLL2* and *EHMT1*, both encoding histone methyltransferases, were identified in Kabuki syndrome and Kleefstra syndrome, respectively.[29,83] The histone acetyltransferases CREBBP and EP300 are implicated in Rubinstein-Taybi syndrome.[158,159] CREBBP is regulated by the ERK/CREB pathway, which is connected to the above mentioned Ras-MAP kinase pathway by the ribosomal protein S6 serine/threonine kinase RPS6KA3. Mutations in this gene cause Coffin-Lowry syndrome.[160] Only recently, mutations in several components of the SWI/SNF ATP-dependent chromatin remodeling complex were identified in Coffin-Siris syndrome, Nicolaides-Baraitser syndrome, and unspecific ID.[76,117,161-164] This complex regulates gene expression by using energy of ATP hydrolysis to alter chromatin structure around its target genes in order to facilitate access of other transcription factors.[165] Other chromatin remodeling ATPases/helicases are for example ATRX in X-linked mental retardation with alpha-thalassemia[166] and CHD7 in CHARGE syndrome.[36]

1.6.6. Cytoskeleton

Proper regulation of the actin and microtubule cytoskeleton is important for dendritic spine morphology and synaptic activity.[4] Recently, mutations in the cytoplasmic actin genes *ACTB* and *ACTG1* were identified in patients with Baraitser-Winter syndrome.[167]

1.6.7. Channelopathies

Most of the so far known channelopathies are implicated in idiopathic epilepsy or epilepsy syndromes.[168] Many of them are associated with developmental and cognitive anomalies. Mutations in sodium channel genes like *SCN1A* and *SCN2A* as well as in the potassium channel gene *PCDH19* cause a broad range of epileptic and neurodevelopmental disorders ranging from familial febrile seizures[169] to severe and early-onset epileptic encephalopathies including Dravet syndrome.[170,171]

1.6.8. Ciliopathies

Cilia are complex sensory organelles. They play a role in the regulation and control of various cellular signaling pathways. Ciliopathies are a heterogeneous group of disorders which are caused by mutations in genes/proteins that localize to the cilium-centrosome complex.[172] Apart from renal disease and retinal blindness, ciliopathies comprise also Joubert and Bardet-Biedl syndromes. Both disorders are accompanied by brain malformations, neurological symptoms and cognitive impairment. Therefore, a role of the

primary cilium and of the many genes involved in its regulation is assumed for the modulation of neurogenesis, neuronal migration, cell polarity, axonal guidance, and possibly adult neuronal function.[173,174] Both Joubert syndrome and Bardet-Biedl syndrome are genetically heterogeneous disorders with mutations in several ciliary genes resulting in an highly similar phenotype, respectively.[175,176]

1.7. Modules and networks

On one hand a distinct ID disorder can be caused by mutations in different genes, whereas on the other hand mutations in one particular gene can result in several clinically distinct phenotypes. This general complexity in disease genetics has for a long time being reflected in the so called “lumping” and “splitting” debate of phenotypes.[177] Together with the growing number of identified ID related genes, the awareness of functional connection between those is increasing. It has been observed that similar phenotypes are often caused by defects in functionally related genes.[178] The so far best studied “module” are the above already mentioned “Rasopathies”, comprising the Noonan/CFC/Costello syndrome spectrum and being caused by defects in several proteins of the Ras-MAPK pathway.[179] Also other correlations between similar clinical phenotypes and their underlying molecular modules are reflected in comprehensive terms such as “ciliopathies”, including ID disorders like Joubert and Bardet-Biedl syndromes, “cohesinopathies”, comprising Cornelia-de-Lange syndrome and related disorders, and “channelopathies”.[180]

These terms are based on the attempt to express molecular themes behind clinical phenotypes. This is accomplished on different levels, emphasizing organelles, processes, pathways or complexes.

Establishing modules and networks is considered to be an important step to be able to cope with the large amount of data that will result from NGS, hence for evaluating and prioritizing variants found in a large number of candidate genes.[96,180] It is also considered to be the starting point for the development of therapies in ID disorders as intervention into a central point of a larger complex or pathway is the only feasible approach compared to targeting every single rare genetic defect. Modules, networks and common themes in ID are subject of chapter 6 of this thesis.

Large scale approaches to establish both clinical and functional networks are currently mainly based on bioinformatic analyses. These can link up information such as protein-protein interaction, known associations with pathways, and phenotype ontology databases.[96,178,181] First approaches in this direction were recently undertaken for autism, autism spectrum disorders, and ID. They indeed showed common pathways of overlapping synaptic regulatory subnetworks, beta-catenin/chromatin remodeling networks, or enrichment

of neurophenotypes.[47,96,182] However, these bioinformatic analyses are hypothetical and rely on the existence, extent and availability of molecular and functional data of the genes and proteins of interest.

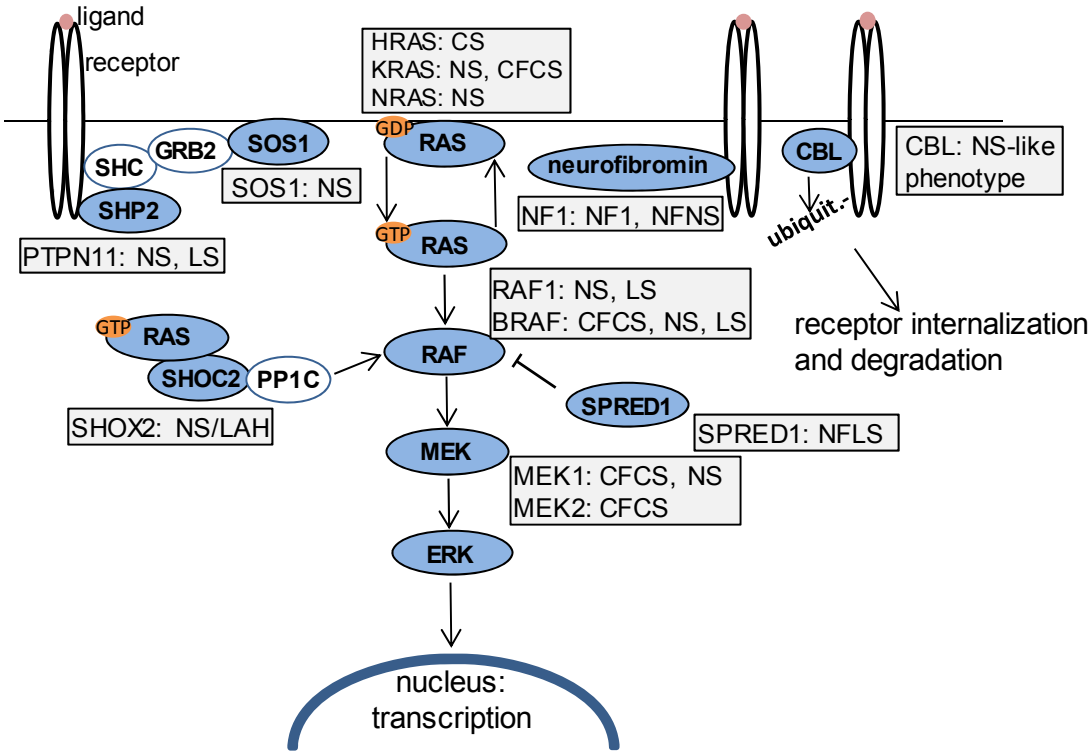

Figure 3 Schematic drawing of the RAS-MAPK signal transduction pathway, one of the best characterized disease-associated pathways.

Dimerized cell surface receptors are activated by ligands, and signals are conducted via the pathway to the nucleus. Shaded ovals represent genes/proteins associated with disorders from the neuro-cardio-facial-cutaneous syndrome family. Light grey boxes name the disorders linked to mutations in the particular gene. NS, Noonan syndrome; CS, Costello syndrome; LS, LEOPARD syndrome; CFCS, cardio-facio-cutaneous syndrome; NF1, neurofibromatosis type 1; NFLS, neurofibromatosis type 1-like syndrome (Legius syndrome); NFNS, neurofibromatosis-Noonan syndrome; NS/LAH, Noonan-like syndrome with loose anagen hair. The figure is adapted from Tartaglia et al..[183]

1.8. Taking the next step: functional analyses

1.8.5. Aims

Functional analyses of ID related genes or proteins are done for different reasons and with different aims. These include: 1) To prove pathogenicity of a specific mutation. This might be either the first one in a newly identified candidate gene and therefore lacking affirmation through a larger number of patients. Or it might be found in a known disease gene but is in

its nature or localization different from previously known mutations. 2) To delineate genotype phenotype correlations. Understanding specific effects of specific genetic aberrations helps to appreciate the complex genetic and phenotypic interactions. 3) To find out more about the gene/protein of interest and its function, in order to confirm its role in disease-relevant phenotypes and to generate fundamental knowledge about it. 4) To establish the position and role of a particular gene or protein in a larger complex or pathway. This might help to identify new candidate genes for cognitive function and dysfunction due to their functional and interactive role. Furthermore, it is also an important prerequisite for the establishment of therapeutic approaches that can target a pathway or complex rather than each of its single components.

1.8.6. Sources and approaches to investigate gene/protein function

Depending on the nature of the investigated gene or protein, on the nature of the identified defect, and depending on available patient material, a large variety of functional analyses can be exploited to investigate the general function of a gene/protein and/or to test a more specific outcome of single mutations. Some examples are given below.

a) Human samples

Human samples are usually limited to easily accessible materials like cells from peripheral blood (lymphocytes or cells from EBV-transformed lymphoblastoid cell lines) or fibroblasts. Rarely, brain specimens are available from autopsies or after brain surgery of patients.[184] Usually, these allow only static biochemical or morphological examinations.[185]

Analyses done on peripheral cells from blood or skin do not necessarily reflect the processes happening in brain, and they do not allow investigations on specific neuronal cell types. Therefore, scientists are trying to overcome these limitations by using stem cells that can be differentiated into various subsets of neurons. Utilizing human embryonic stem cells as done for example to model chromosomal disorders[186] was in that case limited to the availability due to spontaneous abortion of aneuploid embryos. It is in general under ethical discussion and prohibited in many countries. However, the derivation of induced pluripotent stem cells (iPS cells) from somatic tissues such as fibroblasts can provide a good source of patient-specific stem cells. It is a very promising perspective for disease research that allows cellular modeling of disease processes,[187] and has already been used successfully in research into Rett syndrome and related disorders.[188,189] However, to what extent these cells mirror the state of (specific) brain neurons still has to be elucidated in more detail.

b) Cell system based analyses

If no suitable patient material is available or if more basic processes are to be investigated, standard cell lines such as HEK (human embryonic kidney) or HELA (cervical cancer) can be

exploited to investigate disease gene function. Overexpression or knockdown of a gene of interest can be induced, respectively. Also a large number of more specific assays are possible when transfecting cell lines with wild type or mutant constructs. Subsequently, intracellular changes of localization can be determined, or transcriptional activity and cellular processes can be measured by image-based approaches or in cell extracts. The latter is described below in more detail.

c) Animal models

Animal models can be used to model pathologies leading to ID and to investigate the role of specific alterations more detailed than usually possible in humans or in human samples. They allow phenotyping on many different levels, thus connecting molecular, morphological, and functional phenotypes with system properties such as changes in learning, memory, and behavior.

Genetically modified mice are currently most commonly used. Mice are, among mammals, the most amenable to genetic manipulation. Furthermore, an extensive knowledge of the murine genome, physiology, and behavior facilitates interpretation of effects.[190] Standardized procedures such as knocking out genes by homologous recombination or random insertion of wild-type or mutant transgenes are used. These may induce pathological events that mimic the human disorder and can therefore provide a basis to study the molecular basis of specific pathologies. Once validated, they can also be used to test potential therapeutic interventions.[190]

Another powerful animal model for neurodevelopmental disorders is *Drosophila melanogaster* due to its rapid generation time, comparatively low cost, and vast arsenal of genetic and transgenic capabilities.[191] Despite the evolutionary and neuro-anatomical divergence between flies and humans, a large number of genes known to be involved in ID have orthologs in *Drosophila*. This, and much other evidence indicate a conservation of molecular mechanisms underlying learning and memory.[192-195]

In addition to generating information on specific gene function and associated pathologies for several ID disorders,[191,192] *Drosophila* allowed a first pharmacological rescue of certain Fragile X- associated cognitive phenotypes in flies.[196] This resulted in further studies in the Fragile X mouse model and triggered ongoing clinical trials in human patients.[197,198]

While the generation of knockout mutants can be quite time consuming for *Drosophila* and even more for mice, the UAS-GAL4 system in *Drosophila* allows rapid and tissue specific knockdown or overexpression of any gene of interest and therefore provides a suitable foundation for many subsequent experiments.[199,200]

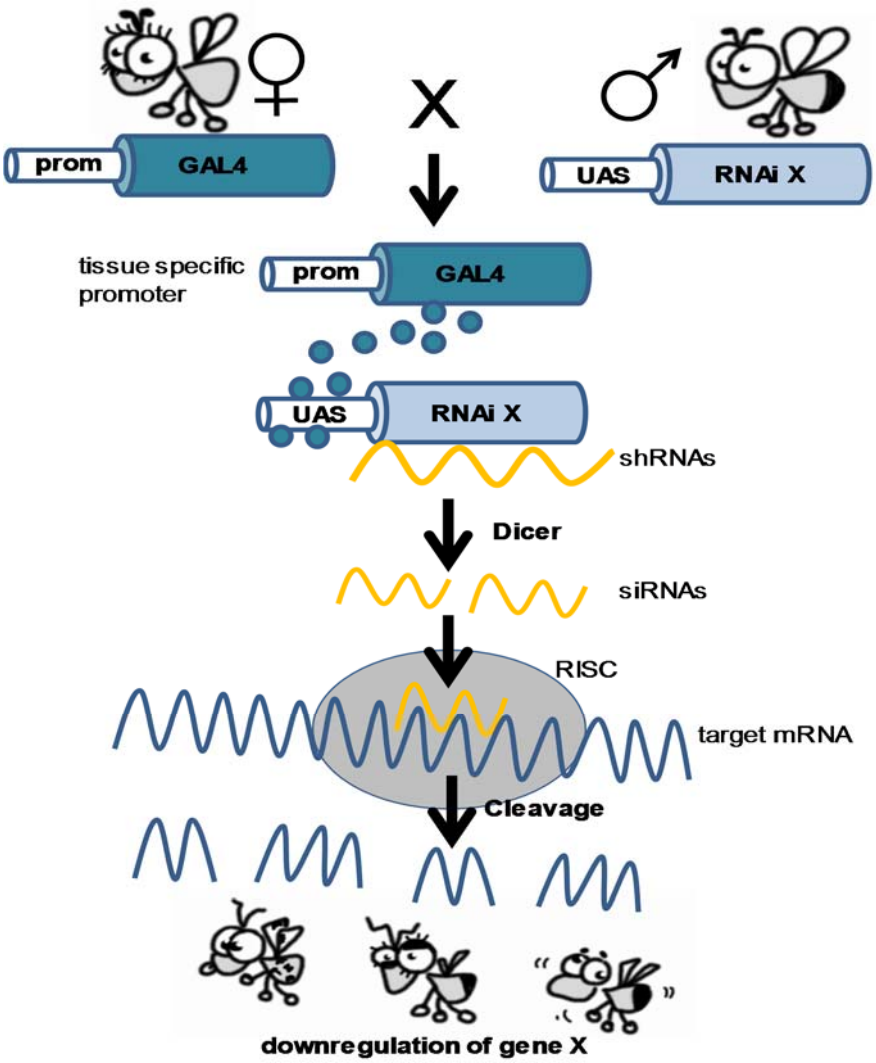

Figure 4 Schematic drawing of RNA interference[201] via the UAS-Gal4 system.
Female, non-mated flies carrying a promoter (prom)-GAL4 insert, are crossed with male flies carrying a UAS-RNAi transgene. In the progenies, GAL4, a transcriptional activator, is generated in specific tissues and binds to the upstream activating sequence (UAS). This activates the RNAi insert to express short hairpin RNAs. These are cleaved by the enzyme Dicer into short interference RNAs (siRNAs), which assemble with the target mRNA via the RNA-induced silencing complex (RISC). The subsequently cleaved mRNA cannot be translated anymore, and therefore a downregulation of the gene of interest is induced.

1.8.7. Methods to investigate gene/protein function

a) In silico analyses

When trying to evaluate the pathogenic potential of an unknown variant, software tools that consider effects on splicing, conservation, or nature of the amino acid change are often used. These provide a first step of evaluation and can be of great assistance in prioritizing genetic variants. However, those *in silico* tools are limited to a theoretical assessment of the variant

based on several parameters. They lack biological evidence that can only be obtained by experimental analyses.

b) RNA

If a genetic defect is assumed to affect a dosage sensitive gene, the expression level of this gene in RNA from the patient can be tested. In order to confirm dosage loss due to a disruption of the gene, expression levels can be determined by quantitative real time-PCR approaches.[202] Quantitative RT-PCR can also be used to investigate effects of a mutation on the expression level of a small number of suspected target genes.[75]

Additionally, individual genetic defects can contribute to the understanding of gene and protein function on a more general level. With the evolving technologies of microarray analysis and transcriptome mRNA sequencing (RNA-Seq) it is now possible to reveal effects of mutations or gene dosage on the whole transcriptome. This facilitates the identification of pathways and networks that are connected with ID disorders or even with sub-phenotypes as shown for example for Rett- and Williams-Beuren syndromes.[203,204]

While the availability of patient material is a prerequisite for such analyses, RNA interference using siRNA (small interfering RNA) or shRNA (short hairpin RNA) permits knockdown of gene expression in cellular systems with subsequent transcriptome profiling as done for example for MYST4 in HEK and HELA cells[205] or for Sox2 and Chd7 in murine neuronal stem cells.[206]

c) Protein

Provided that a suitable antibody is available or can be generated, and depending on its binding sequence, protein truncation and reduction or lack of protein level can be detected by western blot analysis.[207] Both wild type and mutant protein localization and co-localization as well as cellular morphology can be investigated by immunofluorescence on patient/control cells or after transfecting cell lines with expression constructs. For example, aberrant actin organization was shown for ACTB and ACTG1 mutations in cell lines of patients with Baraitser-Winter syndrome.[167]

Effects of mutations in enzymes can be tested by specific activity assays as for example done for the serine/threonine kinase RPS6KA3, which is mutated in Coffin-Lowry syndrome.[207]

The effect of mutations in ion channel encoding genes such as *SCN1A* or *SCN2A* can be tested by using patch clamp techniques in transfected cell lines or xenopus oocytes.[208,209]

d) Protein-DNA binding

Transcription factors regulate transcription of target genes by binding to specific sequences in promoter regions and by activating or repressing mechanisms.

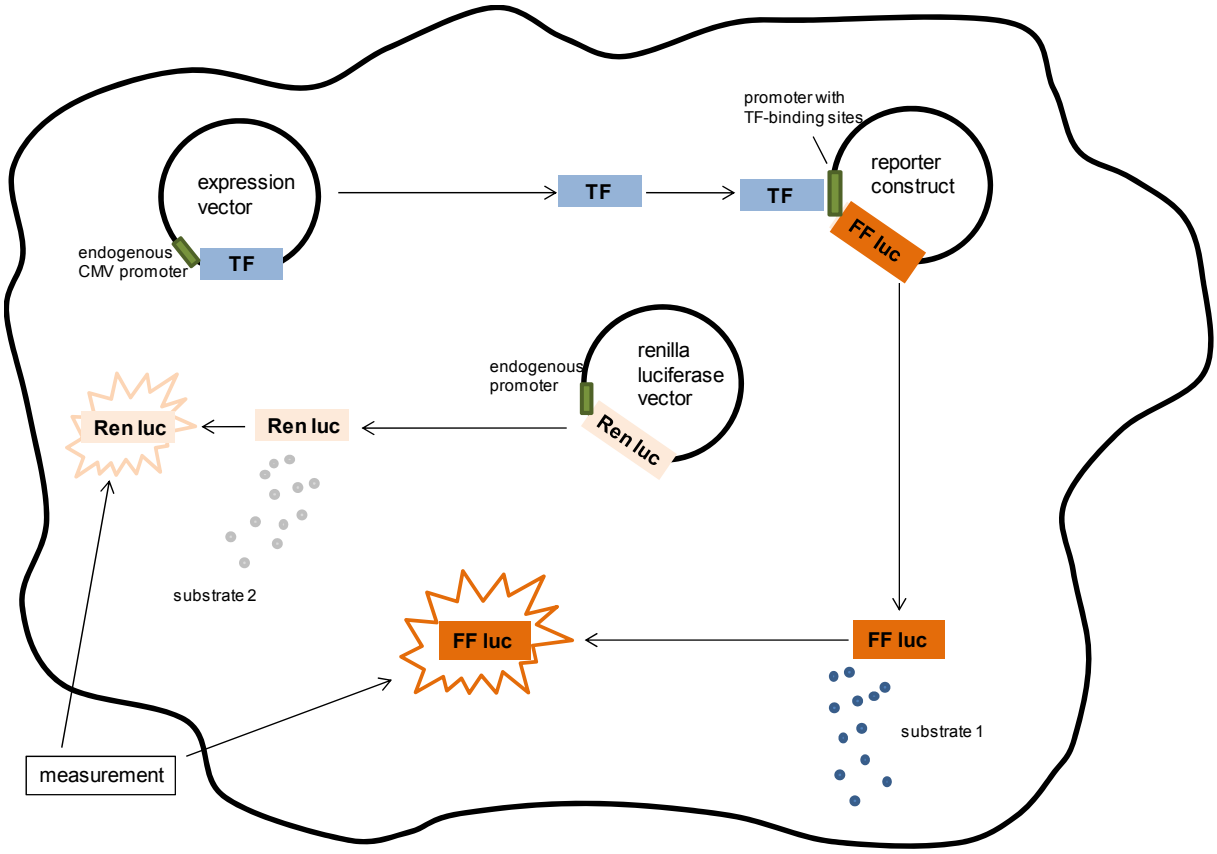

Figure 5 Schematic illustration of an indirect dual luciferase assay.
A cell line is co-transfected with: 1. An expression vector containing wild type or mutant cDNA of the transcription factor (TF) of interest, which is cloned behind an endogenously active promoter. 2. A reporter construct containing cDNA of firefly luciferase (FF luc) and a promoter with specific binding sites of the transcription factor of interest. 3. A construct containing renilla luciferase together with an endogenously active promoter. The transcription factor activates expression of firefly luciferase which is stimulated to produce luminescence by substrate 1. The luminescence is measured and normalized to luminescence intensity of an endogenously active renilla luciferase, in order to even out differences in transfection efficiciency.

Indirect Luciferase assays provide the possibility to test transcriptional activity of a protein by co-transfecting it with a reporter construct. The latter contains the specific binding site of the transcription factor together with a promoter element that activates transcription of luciferase, whose activity can then be determined.[210]

While Luciferase assays allow more complex determination of transcriptional activity, electro mobility shift assays (EMSA) are focusing rather strictly on the pure DNA-binding capacities of a transcription factor. They determine the shift caused by the binding of a protein to a DNA nucleotide during gel electrophoresis.[211] Often both methods are used together in a complementary fashion.[212]

Also chromatin remodeling factors play a role in transcriptional regulation by different mechanisms which regulate chromatin structure and status.[213] The methodology of chromatin immunoprecipitation (ChIP) involves shearing of protein-associated chromatin into smaller fragments followed by immunoprecipitation of the DNA-protein complex with a protein-specific antibody. The isolated DNA-protein complexes can then be dissociated and the specifically enriched DNA segment can be analyzed using PCR amplification methods.[211] ChIP can be combined with subsequent microarrays (ChIP-ChIP) or DNA sequencing (ChIP-Seq), thus allowing generation of genome-wide profiles of binding sites for transcription regulators. Such experiments were done for MYST4 in patient lymphoblastoid cell lines,[205] for Ehmt1 in fly larvae,[213] or for Sox2 and Chd7 in murine neural stem cells.[206]

e) Protein-Protein binding

Co-immunoprecipitation can be used to confirm interaction of a particular protein with the protein of interest by western blot[214] but also to investigate the effect of a mutation on the interaction.[215]

A large number of methods, including co-immunoprecipitation, co-purification, or affinity purification of protein complexes, can be used to identify new interaction partners of the protein of interest. This can be achieved by separating bound proteins on a polyacrylamide gel and by subsequently analyzing them with mass spectrometry.[216] For example, FLAG- affinity based protocols and mass spectrometry were used to identify factors of the nucleosome remodeling and deacetylation (NuRD) complex as interaction partners of Sox 2[206] or of PHF6.[217]

While these methods require generation of homogenized protein extracts, the yeast two-hybrid technique allows detection of interacting proteins in living yeast cells.[216] The classical cDNA-library screen searches for pairwise interactions between the defined proteins of interest (bait) and their interaction partners (preys), present in cDNA libraries or sub-pools of libraries. Interaction between two proteins activates reporter genes that enable growth on specific media or a color reaction.[216] Yeast two-hybrid approaches were for example used to identify or confirm interaction partners of Fragile X syndrome protein FMR1,[218] of CASK, which is implicated in X-linked ID,[219] or for Joubert syndrome related proteins within the ciliary network.[220]

f) Phenotyping in animal models

Over the last years, model organisms such as mouse and *Drosophila* have contributed to a broader understanding of the underlying molecular and functional mechanisms in a large number of ID disorders. Examples are Down syndrome, Angelman syndrome, Rett syndrome, and Fragile X syndrome.[190-192]

Basic investigation of expression patterns of a gene of interest during different stages of embryonic and adult development can be investigated with whole mount in situ hybridization in both mouse[221] or *Drosophila.*[222]

The nervous system of knockout or knockdown models allows various investigations such as detection of expression levels, investigation of morphology of neuronal and glial cells, or testing synaptic function.

For example, the αPix/Arhgef6 mouse model for X-linked ID allowed to determine expression of the gene in certain areas of the brain and to detect altered expression of Rho GTPases in the hippocampus. Furthermore, changes in dendritic morphology and linear spine density as well as altered synaptic plasticity as measured by long-term potentiation and repression, were observed.[223] Similarly, the Ehmt1 knockout fly allowed delineation of expression patterns in the brain and detection of altered dendrite development in multidendrite neurons.[213]

Additionally, deficits in navigation, object exploration and complex positional learning were observed in the αPix/Arhgef6 knockout mouse.[223] In the Ehmt1 knockout mouse model, reduced activity and exploration increased anxiety, diminished social playing, and a delayed or absent response to social novelty were demonstrated.[224] The Ehmt1 mutant fly showed deficiencies in larval locomotor behavior, habituation and courtship memory.[213]

Using the UAS-GAL4 system the consequences of a knockdown in specific tissues can be tested. This was done for climbing behavior in the *Drosophila* model of CHARGE syndrome after inducing knockdown of kismet/CHD7 in different subsets of neurons and muscle cells.[225] Courtship memory defects in EHMT1 flies have been show to reside in mushroom bodies.[213]

These examples illustrate that a vast number of behavioral tests in mice and *Drosophila* exist. Mouse models can mirror human disease symptoms like autistic traits quite well in behavioral assays, which test impaired social interaction, communication deficits, and repetitive behaviors.[226] However, in general no one-to-one equivalency of human behavioral phenotypes to animal models can be expected. Modeling of human-like symptoms in animals should therefore not test if the animal would show a given cognitive impairment but rather, how a cognitive impairment would manifest in the animal.[185]

1.9. Outline and aim of this thesis

The aim of this thesis was to contribute to the genetic, clinical and functional characterization of ID disorders, with a particular focus on Pitt-Hopkins (PTHS) and Pitt-Hopkins-like syndromes.

Work described in this thesis covers the identification of *TCF4* haploinsufficiency as the underlying genetic defect for Pitt-Hopkins syndrome, a distinct, but previously unrecognized syndrome that is characterized by severe ID, distinct facial features, and breathing anomalies. Consequences of TCF4 mutations on interaction with ASCL1 from the ASCL1-PHOX2B-Ret pathway were investigated with a transcriptional reporter assay as those might explain some of the characteristic PTHS symptoms (**chapter 2**).

Following the identification of the underlying genetic defect, a larger number of patients with Pitt-Hopkins syndrome and *TCF4* mutation were collected in order to further delineate the clinical and genetic characteristics of this syndrome (**chapter 3**).

As in many patients with a relatively homogenous PTHS-like phenotype no *TCF4* mutation was identified, we performed molecular karyotyping in order to identify further defects responsible for PTHS or related disorders. We found a homozygous deletion in *CNTNAP2* in a pair of siblings and a compound heterozygous deletion and splice site mutation in *NRXN1* in a sporadic patient. Subsequently, *Drosophila melanogaster* was used as a model organism to investigate a possible interaction between Nrx-IV and Nrx-I that could provide an explanation for the similar phenotypes caused by the newly identified recessive defects in their human homologues *CNTNAP2* and *NRXN1* (**chapter 4**).

To further characterize the phenotype associated with defects in *CNTNAP2* and *NRXN1* and to further confirm their role in severe ID, we identified a larger number of patients with CNVs or mutations in either gene and described their genotypic and phenotypic characteristics (**chapter 5**).

Beyond these focused studies, this thesis also aimed at the contribution to a more global understanding of the molecular pathology of ID disorders. In order to pave the way to large-scale functional studies, a systematic inventory of all ID-related genes and their associated phenotypes was established. We also classified ID disorders according to their clinical phenotypes and attempted to establish correlations between clinical and molecular/functional aspects (**chapter 6**).

A general discussion of the findings in this thesis and their implications in the broader field of ID genetics is provided in **chapter 7**.

Chapter 2

Haploinsufficiency of *TCF4* causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome).

Zweier C, Peippo MM, Hoyer J, Sousa S, Bottani A, Clayton-Smith J, Reardon W, Saraiva J, Cabral A, Goehring I, Devriendt K, de Ravel T, Bijlsma EK, Hennekam RC, Orrico A, Cohen M, Dreweke A, Reis A, Nürnberg P, Rauch A.

Am J Hum Genet. 2007 May;80(5):994-1001

Pitt-Hopkins syndrome is a rarely reported syndrome of so-far-unknown etiology characterized by mental retardation, wide mouth, and intermittent hyperventilation. By molecular karyotyping with GeneChip Human Mapping 100K SNP arrays, we detected a 1.2-Mb deletion on 18q21.2 in one patient. Sequencing of the *TCF4* transcription factor gene, which is contained in the deletion region, in 30 patients with significant phenotypic overlap revealed heterozygous stop, splice, and missense mutations in five further patients with severe mental retardation and remarkable facial resemblance. Thus, we establish the Pitt-Hopkins syndrome as a distinct but probably heterogeneous entity caused by autosomal dominant de novo mutations in *TCF4*. Because of its phenotypic overlap, Pitt-Hopkins syndrome evolves as an important differential diagnosis to Angelman and Rett syndromes. Both null and missense mutations impaired the interaction of TCF4 with ASCL1 from the PHOX-RET pathway in transactivating an E box–containing reporter construct; therefore, hyperventilation and Hirschsprung disease in patients with Pitt-Hopkins syndrome might be explained by altered development of noradrenergic derivatives.

In 1978, Pitt and Hopkins described two patients with sporadic “mental retardation, wide mouth and intermittent overbreathing”[227] Since then, only four other sporadic cases with a similar phenotype and one sib pair with possible Pitt-Hopkins syndrome (PHS) have been published, but there is no MIM entry for this entity.[228-231] Breathing abnormalities in these published patients with PHS appeared in midchildhood, were present only when they were awake, and consisted of abrupt paroxysms of tachypnea followed by breath holding and even overt cyanosis. Other common findings were epilepsy with severe grand mal seizures, short stature, microcephaly, severe motor and mental retardation, and minor brain abnormalities such as cerebellar and vermis hypoplasia, hypoplasia of the corpus callosum, small hippocampus, and bulging caudate nuclei. Facial features were characterized by heavy supraorbital regions, a broad and beaked nose with a high bridge and flaring nostrils, a wide mouth, broad palate, and a bowshaped upper lip.[229]

Since extensive metabolic studies and conventional karyotyping did not reveal any clues regarding the etiology, we performed molecular karyotyping[10] using GeneChip Human Mapping 100K SNP arrays (Affymetrix) in the two sporadic cases published by Peippo et al.[229] This work was performed as part of our research study addressing the genetics of mental retardation, which was approved by the Research Ethics Committee of the Medical Faculty of the University of Erlangen-Nuremberg. For molecular karyotyping, DNA samples were hybridized to GeneChip Human Mapping 50K Xba240 and Hind240 arrays, and images were obtained using an Affymetrix GeneChip Scanner 3000. Raw data were analyzed with the Affymetrix copy-number analysis tool (CNAT 2.0.0.9), with 0.5-Mb sliding windows for the

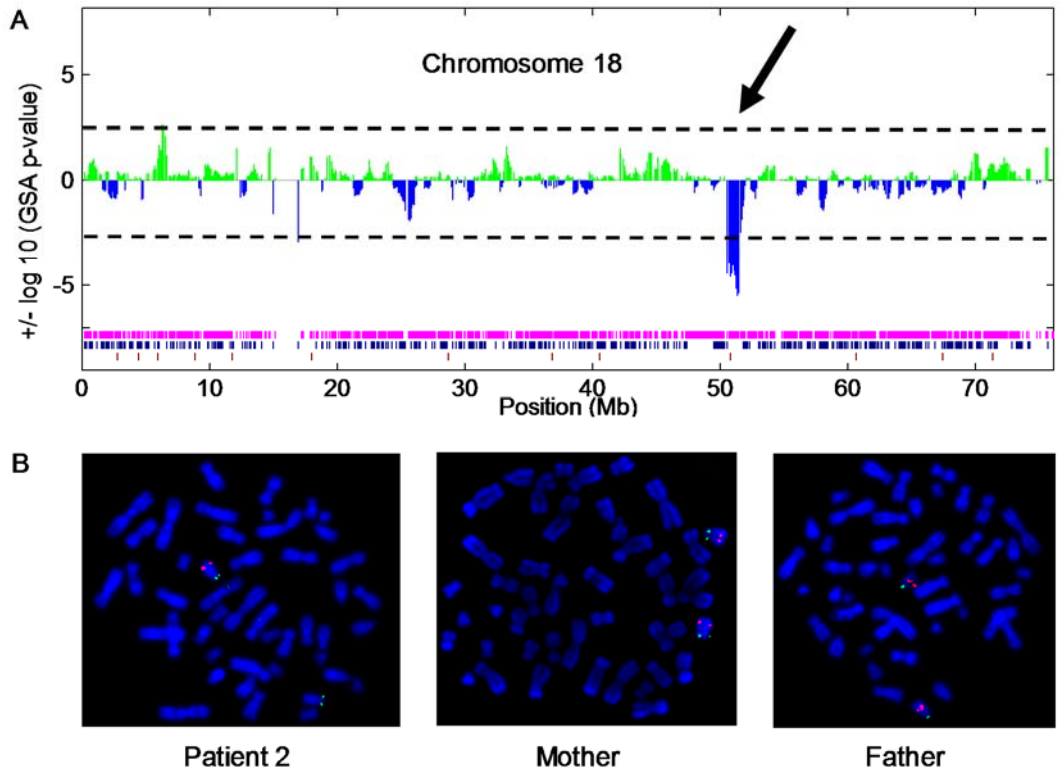

Figure 1 Results of molecular karyotyping with Affymetrix GeneChip Human Mapping 100K SNP array in patient 2.

A) Plot of $\pm\log_{10}$ GSA *P* values for SNPs covering chromosome 18, with use of a Gnuplot program. The 1.2-Mb deletion in 18q21.2 is visible through a cluster of SNPs with GSA *P* values < -2.5. The 63 SNPs indicating a deletion are flanked by SNP_A-1695165 (*rs4800947*; 50,552,638 Mb) and SNP_A-1724163 (*rs784395*; 51,742,365 Mb). Genotypes of individual SNPs are indicated by colored bars at the bottom (magenta=homozygous; blue=heterozygous; brown=no call). Note the magenta stretch of homozygosity corresponding to the deleted region. **B)** Representative results of two-color FISH analyses, with the RP11-7L24 probe labeled with Cy3 (*pink*) in combination with a FluoroX-labeled (*green*) subtelomeric 18p control probe, for patient 2 and his parents. Whereas the RP11-7L24 probe is lacking on one chromosome 18 homologue in the patient, it is present on both homologues in the parents, demonstrating de novo origin of the deletion in patient 2.

genomic smoothing algorithm (GSA). Copy-number values calculated by the CNAT were filtered for clusters of SNPs with high GSA *P* values with use of a self-programmed software tool that we named “CNVFinder” (J. Hoyer, A. Dreweke, C. Becker, I. Göhring, C. Thiel, M. M. Peippo, R. Rauch, M. Hofbeck, U. Trautmann, C. Zweier, M. Zenker, U. Hüffmeier, C. Kraus, A. Ekici, F. Rüschendorf, P. Nürnberg, A. Reis, and A. Rauch, unpublished material). In one of the two patients analyzed, molecular karyotyping revealed a 1.2-Mb deletion on 18q21.2 (fig. 1A). The deletion was confirmed by FISH analysis with 18q21.2 BACs RP11-99A1 (RAB27B; 50.5–50.7 Mb), RP11-839G9 (CCDC68; 50.7–50.9 Mb), and RP11-7L24 (TCF4 [GenBank accession number NM_003199.1]; 51.1–51.24 Mb), all of which lacked the

specific signal on one chromosome 18 homologue (GenBank accession number NT_025028.13) (fig. 1B). All three BACs gave normal FISH results in both parents, thus confirming de novo origin of the deletion in the patient. This deletion contained three known genes: *RAB27B* (a member of the RAS oncogene family), *CCDC68* (coiled-coil domain containing 68), and *TCF4* (transcription factor 4). Since, according to the UCSC Genome Browser database, only *TCF4* is highly expressed in fetal and adult brain, we considered it the most likely candidate gene for the PHS phenotype. Sequence analysis of coding exons 2–19 and intronic flanking regions of *TCF4* in the second patient published by Peippo et al.[229] revealed the missense mutation R576/580W within exon 18, which codes for the helix-loop-helix (HLH) domain of TCF4. Sequencing was performed bidirectionally on an ABI 3730 capillary sequencer (Applied Biosystems) (detailed conditions and primer sequences are available on request). *De novo* origin of the missense mutation was proven by its exclusion in both parents. Paternity was verified, and mistake of probes was excluded by genotyping 14 polymorphic microsatellite markers in the child and both parents (PowerPlex 16 System [Promega]).

Because of phenotypic overlap with the Mowat-Wilson syndrome (MIM 235730), both patients had been tested for mutations in the *ZFHX1B* gene before this study.[229,232] We therefore screened 87 patients for *TCF4* mutations in whom *ZFHX1B* testing had revealed normal results. None of these patients with mental retardation and variable features of the Mowat-Wilson syndrome spectrum, including constipation and Hirschsprung disease (HSCR [MIM 142623]), showed a *TCF4* mutation.

We then sequenced 29 further patients with a more specific phenotypic overlap with PHS—that is, with at least two of the following features: severe mental retardation, breathing anomalies, and PHS-like facial dysmorphism. These patients also included a sib pair and a sporadic case formerly published as having PHS or possible PHS.[228,229,231] The original cases published by Pitt and Hopkins[227] were not available, because one patient died and the other was lost to follow-up (D. Pitt, personal communication). The case published by Singh[230] could not be tracked either. *TCF4* mutational analysis revealed three heterozygous stop mutations and a splice-site mutation in four of the unpublished patients (fig. 2 and table 1). De novo origin was proven in two patients (3 and 4), whereas parents of patients 5 and 6 were not available for testing. We therefore excluded both the R385X and IVS9-1G>C mutations of the latter patients in a total of 180 healthy European control individuals.

Since most of our patients have a deletion or stop or splice-site mutations, haploinsufficiency of *TCF4* is likely to be causative of PHS. The only observed missense mutation affects an evolutionarily conserved amino acid and is located within the basic HLH (bHLH) domain of TCF4, thus likely impairing the binding capacity of the only functional domain known so far.

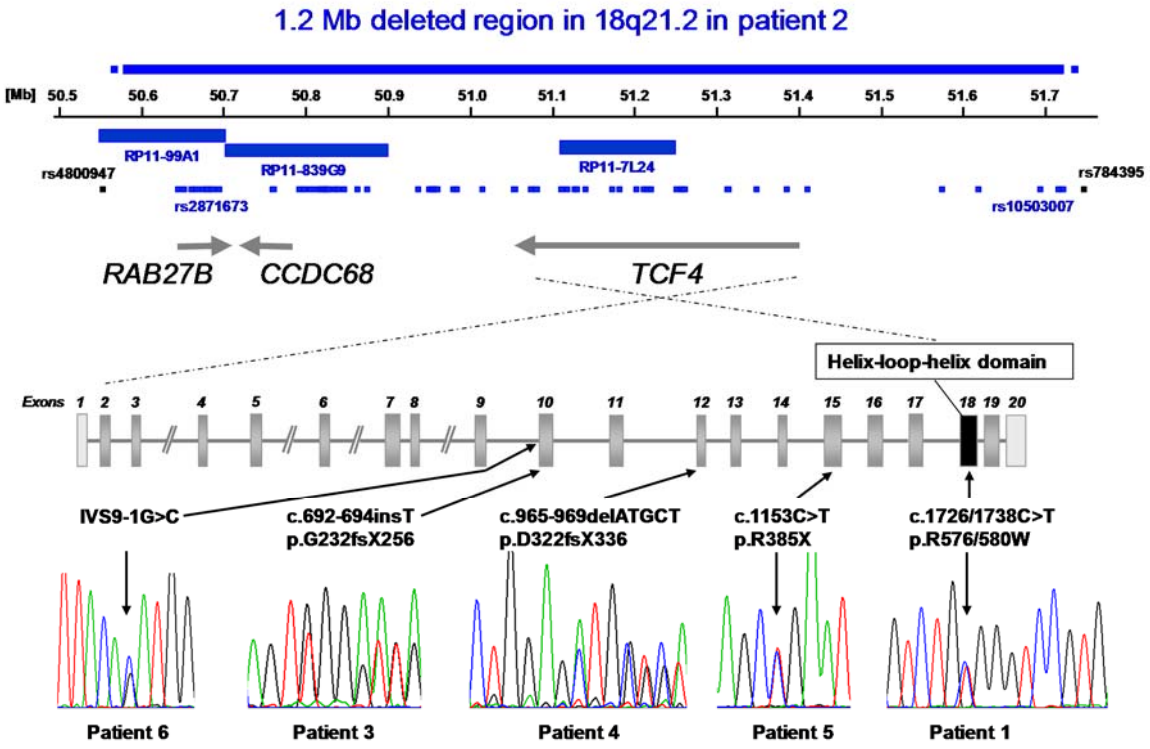

Figure 2 Schematic drawing of the 1.2-Mb deletion region in patient 2 and location and electropherograms of *TCF4* mutations in patients 1, 3, 4, 5, and 6 within a schematic drawing of exon-intron structure of *TCF4.*

Noncoding exons are light gray, and the exon coding for the functional domain is black. Blue bars representing BAC clones and blue squares representing SNPs indicate deletion of respective probes, whereas nondeleted SNPs are depicted as black squares.

TCF4 (also called "ITF2," "E2-2," and "SEF2") consists of 20 exons (exons 1 and 20 are noncoding), spans 360 kb, and encodes at least two isoforms of the transcription factor 4 protein, differing in the presence of 4 aa (RSRS) 17 residues before the HLH domain. TCF4 belongs to the class A subfamily of bHLH transcriptional regulators—also called "E proteins," since their basic domain binds to the E-box motifs 5'-ACANNTGT-3' or 5'-CCANNTGG-3'.[233] E proteins are characterized by a broad expression pattern and the ability to form both homo- and heterodimers with other classes of HLH proteins that are tissue specific or lack the basic DNA-binding domain.[234,235]

Homozygous *Tcf4* deletions in mice lead to early lethality of unknown reason and a slight decrease in pro-B cell numbers.[235] In contrast to the apparently normal single heterozygous *Tcf4*-knockout mice, transheterozygous knockout combinations of any two of the E proteins Tcf4, E2a, and Heb generate significantly reduced numbers of pro-B cells, and mice conditionally mutated in *Tcf4* show a partial block in both B and T lymphocyte development.[234,235] Accordingly, none of our patients with heterozygous *TCF4* mutations showed clinical evidence of immunodeficiency, and lymphocyte and Ig counts in patient 3

Table 1 Clinical findings for patients with *TCF4* mutations

Patient	1*	2*	3	4	5	6
Sex	M	M	M	M	F	M
Age (years)	14	11	8	12	29	29
Birth Weight Length OFC	2800 g (>P3) 48 cm (P10) 33 cm (P3)	2620 g (>P10) 47 cm (P10) 33 cm (P25)	2520 g (>P25) 49 cm (>P50) 32.5 cm (P25)	2450 g (P10) 46 cm (P3) 31.5 cm (P10)	3660 g (P50) ?53 cm (<P75) NA	2500 g (P3) NA NA
Postnatal Weight Height OFC	NA -3SD -3SD	NA -3SD -5SD	Weight P75 P10-25 P25	Weight P10 P25 P25	Weight <P3 <P3 <P3	Weight <P3 <P3 <P3
Walking	Steps with aid	Steps with aid	Steps with aid	5y, ataxic	5y, ataxic	7y, wide based
Speech	No	No	Single words	No	No	No
Seizures	9y	8y	No	No	No	No
Hypotonia	Yes	Severe	Severe	Severe	Yes	Yes
H-A Age of onset	Daytime 5y	Daytime 6y	No	Daytime 8y	Yes 2y	Yes 5y
MRI anomalies	HCC, small hippocampus, BCN	HCC, small hippocampus, BCN	Dilated cerebral ventricles	HCC, BCN	NA	CT only normal
SDFC	No	No	Yes (3, 4, bil.)	Yes (2, 3, 4 bil.)	Yes (3, 4 bil.)	Yes (4)
SPC	Yes	No	Yes	Yes	Yes	NA
Happy disposition	Yes	Yes	Yes	Yes	unmotivated laughter episodes	Yes
GI-anomalies	Constipation	HSCR	No	No	Severe constipation	No
Other anomalies	Broad finger tips, wide and flat palate	Broad finger tips, wide and flat palate	Strabism, scoliosis, sacral dimple, fetal finger pads, overriding 5th toe, short metatarsal V	Supernumerary nipple, scoliosis, long and slender fingers and toes	Strabism, very anxious, autoaggressive behavior, short hands and feet, hyperconvex nails	Lymphoma, fetal finger pads, slender feet with short IV/V metatarsals Dislocated hips Frontal cowlick
Testing: AS-Meth. UBE3A	NA NA	NA NA	N N	N N	N NA	NA NA
TCF4	c.1738C>T p.R580W *de novo*	1.2 Mb deletion *de novo*	c.692-694insT p.G232fsX256 *de novo*	c.965-969delATGCT p.D322fsX336 *de novo*	c.1153C>T p.R385X	IVS9-1G>C

All patients have distinct PHS facial features and severe mental retardation. Results of karyotype tests and MECP2 tests were normal for all patients. y, years; AS-meth., methylations-sensitive PCR at the SNRPN locus concerning Angelman syndrome; bil., bilateral; BNC, bulging caudate nuclei; GI-anomalies, gastrointestinal anomalies; H-A, episodes of hyperventilation-apnea; HCC, hypoplasia of corpus callosum; NA, not available; SDFC, supernumerary digital flexion crease; SPC, single palmar crease; P, centile, OFC, occipitofrontal circumference; y, years; * Patients 1 and 2 were published as cases 1 and 2, respectively, by Peippo et al. [229]

showed normal results. Nevertheless, patient 6 developed a Hodgkin lymphoma at age 29 years, which might indicate some kind of regulatory influence of TCF4 on lymphocyte growth. The diversified phenotypes observed with E-protein deficiencies are consistent with the idea that E proteins are involved in dimeric interactions with many different tissue-specific HLH proteins.[235] Ascl1, the mouse homologue of one of these tissue-specific proteins, is highly expressed in specific regions of the developing CNS and in sympathetic and enteric precursor cells, and Ascl1-null mice die at birth.[236] ASCL1 was shown to form complexes with TCF4 that have the ability to bind an E-box.[236] The interaction between TCF4 and ASCL1 is an interesting observation, since mutations in *ASCL1* have been shown to impair the noradrenergic neuronal development in brain stem, causing some cases of congenital central hypoventilation syndrome (CCHS, or Ondine curse [MIM 209880]).[237] In *Ascl1*-knockout mice, an impaired c-RET expression in brain-stem noradrenergic neurons and an increased baseline breathing frequency were reported.[238] The major gene that causes CCHS, *PHOX2B*, also belongs to the RET-signaling pathway, which, in mice, involves the sequential expression of the *Ascl1*, *Phox*, *Ret*, and *TH* genes that are responsible for the development of all transient or permanent noradrenergic derivatives.[237]

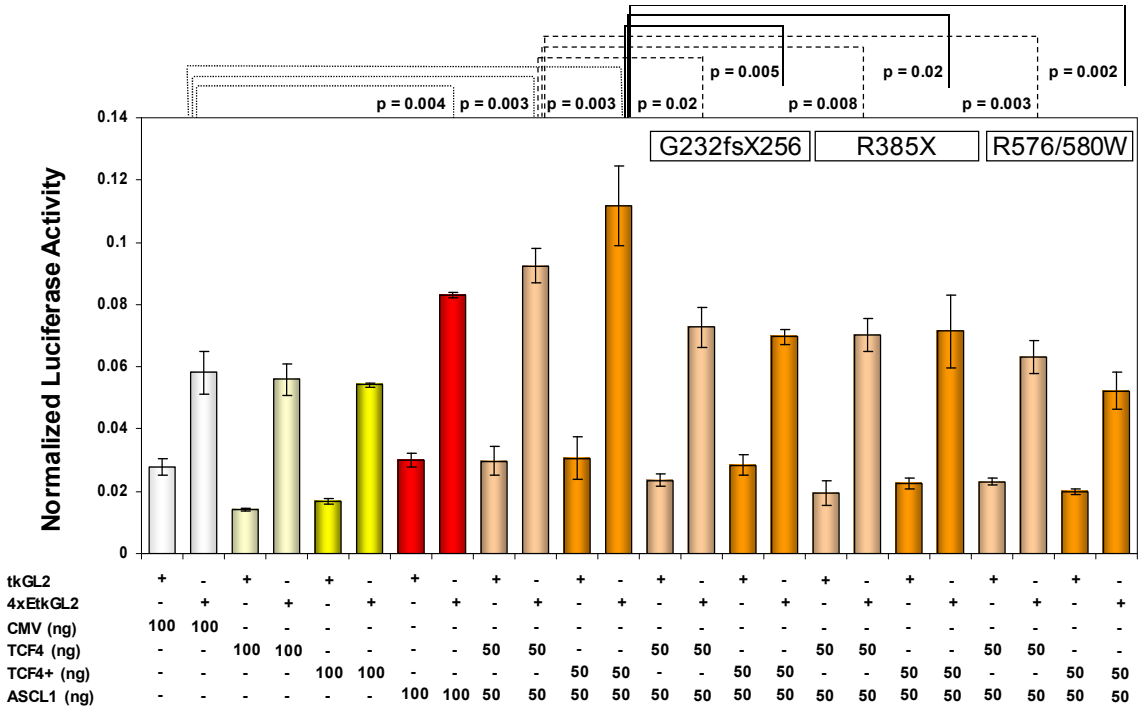

Figure 3 Transcriptional reporter assay showing impaired interaction of TCF4/TCF4^{+} mutants with ASCL1.

JEG-3 cells were transiently transfected with a luciferase reporter construct with a herpes simplex thymidine kinase promoter either without binding sites (tkGL2) or with four E boxes (4xEtkGL2) located within the pTa enhancer. Co-transfection was performed with an empty CMV-expression vector (*white*)

or CMV-expression vectors containing the complete cDNA of either ASCL1 (*red*), TCF4 (*pale yellow*), its splice variant TCF4^{+} (including amino acid RSRS) (*bright yellow*), as well as wild-type or three different mutants of TCF4 and TCF4^{+}, transfected in combination with ASCL1 (*different shades of orange*). Results were normalized for transfection efficiency to a cotransfected renilla luciferase vector and were expressed as mean values with SD of three independent transfections. Probably because of endogenous E proteins, cells transfected with the empty CMV vector alone already showed a slight transactivation of the luciferase vector containing the four E boxes (4xEtkGL2), in comparison with the reporter vector without E boxes (tkGL2). TCF4 and TCF4^{+} alone did not increase the activation of the reporter construct but enhanced the observed activation by ASCL1 when cotransfected with the latter. In contrast, TCF4 and TCF4^{+} mutants containing the mutation p.G232fsX256, R385X, or R576/580W did not enhance the activation by ASCL1. Differences in activation levels showed significant *P* values obtained by the Student *t* test.

To investigate this interaction between TCF4 mutants and ASCL1, we established a transcriptional reporter assay using a luciferase reporter construct with a herpes simplex thymidine kinase promoter, either without binding sites (tkGL2) or with four E boxes (4xEtkGL2) located within the pTa enhancer, which interacts with several E proteins.[239] JEG-3 cells (derived from human choriocarcinoma, American-type culture collection cell line HTB-36) were transiently cotransfected with a cytomegalovirus (CMV)–expression vector containing ASCL1, TCF4, its splice variant TCF4+ (including amino acid RSRS), either alone or in combination, and three different mutants of TCF4 and TCF4+ in combination with ASCL1 (fig. 3). Results were normalized, for transfection efficiency, to a cotransfected renilla luciferase vector.

In accordance with the results observed by Persson et al.,[236] who used a transcriptional reporter construct containing four E boxes from the muscle creatine kinase enhancer, in our assay, TCF4 and TCF4+ alone did not increase the activation of the reporter construct but enhanced the observed activation by ASCL1 when cotransfected with the latter. In contrast, TCF4 and TCF4+ mutants containing the mutations p.G232fsX256, R385X, or R576/580W did not enhance the activation by ASCL1 (fig. 3). Therefore, breathing anomalies in patients with PHS may also be caused by impaired noradrenergic neuronal development through defective TCF4 interaction with the ASCL1-PHOX-RET pathway. This interaction might also explain the occurrence of HSCR and constipation in patients 1, 2, and 5, since *RET* is the major gene for isolated HSCR[240] and since patients with CCHS show an increased incidence of HSCR (20%).[241] However, no obvious defect in the sympathetic nervous system in *Tcf4*-knockout mice was observed that could be attributed to a functional interaction with Ascl1.[236] In contrast to patients with CCHS, who are intellectually not impaired, patients with PHS are severely mentally retarded, which may indicate TCF4

involvement in other pathways important for brain development and function corresponding to its high expression in brain.[242]

Figure 4 Facial phenotype of patients with *TCF4* mutations.
A–C) Patient 1 at ages 6 mo (A), 18 mo (B), and 14 years (C). **D** and **H)** Patient 6 at age 29 years. **E–G)** Patient 2 at ages 6 mo (E and F) and 11 years (G). **I–K)** Patient 3 at ages 3 years (I), 6 years (J), and 8.75 years (K). **L** and **M)** Patient 4 at age 12.5 years. Note deep-set eyes; broad and beaked nasal bridge with down-turned, pointed nasal tip and flaring nostrils; wide mouth with widely spaced teeth, and Cupid-bowed upper lip and everted lower lip; mildly cupshaped, fleshy ears; as well as increased coarsening of facial features with age.

All six patients with *TCF4* mutations showed severe mental retardation and striking facial resemblance, at least to patient 1 initially reported by Pitt and Hopkins,[227] consisting of deep-set eyes; broad and beaked nasal bridge with down-turned, pointed nasal tip and flaring nostrils; wide mouth with widely spaced teeth, Cupid-bowed upper lip, and everted lower lip; and mildly cup-shaped, fleshy ears (fig. 4). Further findings were variable, including breathing abnormalities, which were not yet obvious in patient 3 at age 8 years (table 1). Other common signs were magnetic resonance imaging (MRI) anomalies such as hypoplastic

corpus callosum and bulging caudate nuclei, happy disposition or unmotivated laughter episodes, muscular hypotonia, severe constipation or HSCR, single palmar creases, and supernumerary digital flexion creases. The manifestation of a lymphoma in one of the oldest patients may be attributed to the role of TCF4 in lymphocyte development,[234,235] but further observations and TCF4 studies in lymphoma tissues are necessary to address this question.

Since most of the patients in whom *TCF4* mutations were identified had previous testing for Angelman syndrome (MIM 105830) because of facial and behavioral resemblance and for Rett syndrome (MIM 312750) because of late-onset ventilation anomalies and severe mental retardation, PHS might be an important differential diagnosis for these disorders. Interestingly, one patient with severe mental retardation and attacks of deep sighing and hyperventilation who carries an interstitial deletion (18)-(q21.1q22.3) was reported as a case of "atypical Rett syndrome."[243] In contrast, Joubert syndrome (MIM 213300, 608091, 609583, 608629, 610688, and 610188) was never considered in any of these patients, because its characteristic intermittent hyperventilation-apnea anomaly manifests in the newborn. Although most of the patients showed more or less overlapping features, we were unable to demonstrate a mutation in the remaining 24 patients, including the sporadic case published by Van Balkom et al.[231] and the sib pair described by Orrico et al.[228] Although larger gene deletions were excluded by the presence of heterozygous SNPs in all but 2 of the 24 patients, we cannot exclude atypical intronic mutations or single-exon deletions in the 24 mutation-negative patients. Nevertheless, the large number of mutation-negative patients with more or less similar phenotype might indicate the involvement of further genes that possibly interact with TCF4.

Our study shows that molecular karyotyping[10,244,245] (genomewide copy-number profiling) is able not only to disclose novel microdeletion syndromes[23-25]—as well as the underlying gene defect in well-known disorders, as recently shown for the autosomal dominant CHARGE (coloboma, heart anomaly, choanal atresia, retardation, genital and ear anomalies) syndrome (MIM 214800)[36] and the autosomal recessive Peters-Plus syndrome (MIM 261540)[73]—but also to resolve the etiology in very rarely reported phenotypes. However, our findings suggest that PHS is widely underdiagnosed, just as the now clinically recognizable Mowat-Wilson syndrome was until the identification of the underlying gene defect, in 2001.[150,232,246] Notably, PHS is, to our knowledge, the first constitutive phenotype in which the etiology was identified by molecular karyotyping with use of genomewide SNP arrays.

Acknowledgments

We thank the family members for their kind participation, and we thank Daniela Schweitzer, Michaela Kirsch, and Christian Becker for excellent technical assistance. We are in debt to

Franz Rüschendorf, Berlin, for kindly providing his Gnuplot program for visualization of CNAT outputs, and to Christine E. Campbell, Buffalo, for providing the tkGL2 vector. This work was supported by Deutsche Forschungsgemeinschaft grant RA 833/7-1 (to A. Rauch and P.N.).

Web Resources

Accession numbers and URLs for data presented herein are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for TCF4 [accession number NM_003199.1] and chromosome 18 [accession number NT_025028.13])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for Mowat-Wilson syndrome, HSCR, CCHS, Angelman syndrome, Rett syndrome, Joubert syndrome, CHARGE syndrome, and Peters-Plus syndrome)

UCSC Genome Browser, http://genome.ucsc.edu/cgi-bin/hgTracks

Chapter 3

Further delineation of Pitt-Hopkins syndrome: phenotypic and genotypic description of 16 novel patients.

Zweier C, Sticht H, Bijlsma EK, Clayton-Smith J, Boonen SE, Fryer A, Greally MT, Hoffmann L, den Hollander NS, Jongmans M, Kant SG, King MD, Lynch SA, McKee S, Midro AT, Park SM, Ricotti V, Tarantino E, Wessels M, Peippo M, Rauch A.

J Med Genet. 2008 Nov;45(11):738-44.

ABSTRACT

Background: Haploinsufficiency of the gene encoding for transcription factor 4 (TCF4) was recently identified as the underlying cause of Pitt–Hopkins syndrome (PTHS), an underdiagnosed mental-retardation syndrome characterized by a distinct facial gestalt, breathing anomalies and severe mental retardation.

Methods: *TCF4* mutational analysis was performed in 117 patients with PTHS-like features.

Results: In total, 16 novel mutations were identified. All of these proven patients were severely mentally retarded and showed a distinct facial gestalt. In addition, 56% had breathing anomalies, 56% had microcephaly, 38% had seizures and 44% had MRI anomalies.

Conclusion: This study provides further evidence of the mutational and clinical spectrum of PTHS and confirms its important role in the differential diagnosis of severe mental retardation.

INTRODUCTION

In 1978, Pitt and Hopkins described two unrelated patients with sporadic ‘‘mental retardation, wide mouth and intermittent overbreathing’’.[227] After this initial report, only four other sporadic cases and one sibling pair with a similar phenotype of severe mental retardation, wide mouth and breathing anomalies were reported as possible cases of Pitt–Hopkins syndrome (PTHS, OMIM 610954).[228-231] Recently in two patients, two deletions, one of 1.2 Mb and one of 1.8 Mb, in 18q21.1, detected by molecular karyotyping using 100 K single nucleotide polymorphism (SNP) arrays and bacterial artificial chromosome arrays, respectively, led to the identification of haploinsufficiency of *TCF4*, encoding a basic helix–loop–helix (bHLH) transcription factor, as the underlying cause of PTHS.[247,248] A recurrent *de novo* missense mutation in the bHLH domain region of the *TCF4* gene in three patients and one other missense mutation at the same position,[247,248] as well as one splice-site mutation and three stop mutations[248] in further patients were later identified. In the meantime, two other patients with the PTHS or overlapping phenotype and a deletion of *TCF4* were reported.[249,250]

The *TCF4* gene on chromosome 18 encodes a member of the bHLH transcription factor family (also called ‘‘E-proteins’’ as their basic domain binds to the Ephrussi-box (E-box) consensus binding site ‘‘CANNTG’’).[233] These transcription factors are able to bind DNA as homodimers or heterodimers with other classes of HLH proteins and play an important role in many developmental processes, including the differentiation of the vertebrate nervous system and the development of the cortex.[251,252] *TCF4* encodes at least two isoforms, differing in the presence of 4 amino acid residues (RSRS) 17 residues upstream of the HLH domain.[242]

Both null mutations and missense mutations located within the bHLH domain of TCF4 impaired its interaction in vitro with ASCL1, a tissue-specific HLH protein from the PHOX-RET pathway. As this pathway is involved in Hirschsprung disease (HSCR; OMIM 142623) and the Ondine hypoventilation syndrome (OMIM 209880) through its role in the development of noradrenergic derivatives, the finding of HSCR or severe constipation and breathing anomalies in patients with PTHS might be explained by impaired interaction with ASCL1.[248] Investigations of *tcf4* expression in *Danio rerio* embryos showed early expression in the pallium of the telencephalon, the diencephalon, the midbrain tegmentum, the hindbrain and the branchial arches, thus correlating with the phenotypical spectrum in humans with PTHS.[250]

Because of its phenotypical overlap with Angelman and Rett syndromes, we speculated that PTHS might become an important differential diagnosis with these conditions.

To further delineate the genotypic and phenotypic spectrum of PTHS and to establish its frequency in patients with severe mental retardation, we analyzed 117 patients with overlapping clinical features.

METHODS

Ethics approval for this study was obtained from the ethics committee of the Medical Faculty, University of Erlangen-Nuremberg, and informed consent was obtained from the parents or guardians to study the patients.

Patients

Our study population contained 117 patients including two sibling pairs, who were referred to us with severe mental retardation and variable additional features reminiscent of the PTHS spectrum, such as microcephaly, dysmorphic facial gestalt or breathing anomalies. At least 70 of these patients had tested negative for Angelman and/or Rett syndrome.

Molecular testing

DNA samples derived from peripheral blood were screened for *TCF4* mutations by bidirectional direct sequencing of the coding exons 2–19 and the non-coding exon 20 and intronic flanking regions (ABI BigDye Terminator Sequencing Kit V.2.1; Applied Biosystems, Foster City, California, USA) using an automated capillary sequencer (ABI 3730; Applied Biosystems). Primer pairs and PCR conditions are available on request.

Paternity was verified in samples taken from patients 7 and 12 and their parents. Probes were verified by genotyping with 14 polymorphic microsatellite markers (PowerPlex 16 System, Promega, Madison, Wisconsin, USA) to exclude any possibility of mistakes.

Reverse transcriptase (RT)-PCR was performed for patients 12 and 14, using primers located in exons 13 and 17 and exons 5 and 11, respectively, on cDNA obtained from mRNA (Superscript II Reverse Trancriptase; Invitrogen, Carlsbad, California, USA). For patient 12, mRNA was extracted from lymphoblastoid cell lines (RNeasy Mini Kit; Qiagen, Valencia, California, USA) and peripheral blood (PAXgene system; Preanalytix, Franklin Lakes, New Jersey, USA), using commercial kits in accordance with the manufacturers' instructions. Aberrant transcripts were extracted from agarose gel (QiaQuick gel extraction kit; Qiagen) and sequenced after reamplification.

Bioinformatic analyses

Disordered protein segments and linear protein interactions motifs were identified using the software programs DisEMBL[253] and ELM,[254] respectively. The effect of the G358V mutation on the aggregation tendency of TCF4 was assessed using AGGRESCAN,[255] a web-based software program for the prediction of aggregation-prone segments in protein sequences and the analysis of the effect of mutations on aggregation propensities of proteins.

Functional testing

Functional consequences of the G358V mutation were tested with a transcriptional reporter assay as described previously.[248] Immunofluorescence was performed with a primary antibody against TCF4 (ab2233-100; Abcam, Cambridge, Massachusetts, USA) and a CY3-labelled secondary anti-goat antibody (C2821-1ML; Sigma-Aldrich, St Louis, Missouri, USA) on JEG3 cells previously transfected with either wild-type TCF4 or mutant TCF4.

RESULTS

Clinical findings

All 16 patients with proven *TCF4* mutations in this study had severe mental retardation with very little speech (2 patients had < 5 words) or no speech (14 patients) and with limited walking abilities. They resembled each other with a specific facial phenotype characterized by deep-set eyes, broad and often beaked nasal bridge with down-turned, pointed nasal tip and flaring nostrils, wide mouth with widely spaced teeth, cupid's bow upper lip, a protruding lower face, and mildly cup-shaped and fleshy ears (fig. 1). Microcephaly was observed in 56% of the patients and breathing anomalies in 56%. Seizures occurred in 38% of the patients between the age of 0 and 5 years. MRI anomalies such as bulging of caudate nuclei, ventricular asymmetry, agenesis or hypoplasia of the corpus callosum and atrophy of the frontal and parietal cortex were observed in 44% of the patients. Hypotonia and constipation were common findings in 13 and 11 of the patients, respectively. Personality was described as happy in 15 of the patients. Single palmar creases were reported in 11 patients and

scoliosis in 4 patients. Other less common anomalies were myopia and fetal pads in four and six patients, respectively (table 1).

Molecular testing

Sequencing of *TCF4* in 117 patients with severe mental retardation and clinical findings overlapping with PTHS revealed novel mutations in 16 patients. Large deletions of *TCF4* could be excluded by the identification of at least one heterozygous SNP in all but five of the remaining patients. Owing to their facial phenotype, there is a strong suspicion that two of the patients with normal *TCF4* sequencing and exclusion of large deletions of the *TCF4* gene do have PHTS. Single exon deletions or mutations in regulatory elements of *TCF4* and locus heterogeneity may explain these patients and are also possibilities in the other, less characteristic patients.

We identified 12 novel stop mutations, including two splice-site mutations, distributed over the gene, one novel frameshift mutation located towards the C-terminus and resulting in an elongation of the putative protein, two novel missense mutations in exon 18 coding for the HLH domain, and one novel missense mutation in exon 14 (table 2, fig. 2). *De novo* occurrence was proven in all but three stop mutations, for whom parental samples were not available. Assumed probe relationships were confirmed in the patients and both parents in the exceptional missense mutation in exon 14 and the unusual splice-site mutation IVS14+3ARG. The exceptional missense mutation in exon 14 was also excluded in 192 healthy control chromosomes. RT-PCR performed on mRNA from a lymphoblastoid cell line of patient 12 with the IVS14+3ARG mutation revealed an aberrant transcript with skipping of exon 14, which leads to a frameshift that results in a premature stop after three amino acids (fig. 3B). RT-PCR in patient 14, who had the splice-site mutation IVS9+2insGT, also revealed an aberrant transcript, skipping exon 9, which leads to a frameshift resulting in a premature stop after 14 amino acids (fig. 3C).

Bioinformatic analyses

Computational analyses indicated that the N-terminal 550 amino acids of TCF4 are predominantly disordered and do not adopt a globular (domain-like) tertiary structure. There was also no evidence that position 358 is part of a specific linear protein interaction motif that might be destroyed by the mutation G358V (data not shown).

An analysis of the TCF4 aggregation properties revealed that the G358V mutation leads to an increase in aggregation tendency as indicated by the larger hotspot area at the site of the mutation (fig. 3A). Compared with the size of adjacent hotspots, the hotspot emerging at the site of the G358V mutation is more than twice as large.

Functional testing

Functional testing of the missense mutation G358V with a transcriptional reporter assay showed only inconsistently a mild reduction in transcriptional activity compared with the wild type (data not shown). Immunofluorescence with antibodies against TCF4 revealed no visible aggregates (data not shown).

Figure 1 Clinical appearance of a selection of the patients.

Note the characteristic facial phenotype with coarse face, high cheekbones, a beaked nasal tip, a protruding lower face, a wide mouth with cupid's-bow shaped upper lip and wide-spaced teeth. **A)** Patient 3, aged 22 months; **B)** patient 12, aged 35 months; **C)** patient 15, aged 35 months; **D)** patient 4, aged 3 years; **E)** patient 10, aged 10 years; **F)** patient 9, aged 14 years (right) and 16 years (left); **G)** patient 6, aged 17 years; **H)** patient 2, aged 17 years; **I)** patient 14, aged 18 years; **J)** patient 7, aged 20 years; **K)** patient 5, aged 20 years; **L)** MRI of patient 12 (note the bulging of the caudate nuclei); **M)** hand of patient 12.

DISCUSSION

Clinical spectrum

Including our 16 novel patients, 27 patients with molecularly confirmed PTHS are currently known.[247,248,250] All of these patients are severely mentally retarded with no or only very limited speech and limited mobility. The earliest reported walking age was 2 years in patient 8, whereas some of the older patients cannot walk without support (e.g. patient 2) or not at all (e.g. patient 1). Gait is often unstable and ataxic. Furthermore, muscular hypotonia seems to be a common feature particularly at younger ages.

Breathing anomalies occurred in 18 of 27 (67%) patients, with age of onset varying from a few months to teenage years. These episodes were characterized by daytime periods of hyperventilation followed by apnoea. Milder anomalies may also occur such as ‘‘playing with breath’’ without apnoea in patient 9 or a singular occurrence of hyperventilation after narcosis in patient 16. The oldest patients without breathing anomalies were 18 years old.
Microcephaly, both congenital and acquired, occurred in about 63% of all known patients. While birth weight was in the lower normal range in the patients in our first study,[248] in this cohort birth weight was normal or high normal in all cases for whom this information was available.

Further common symptoms were seizures with onset from birth up to 9 years of age (44%), constipation (67%) and minor anomalies such as single palmar creases and supernumerary phalangeal flexion creases (63%). Hirschsprung disease, which had occurred in one of the first patients,[248] was not noticed in any of the present cohort.

In 14 of 22 (64%) patients who had MRI studies performed, anomalies such as bulging of caudate nuclei, ventricular asymmetry, agenesis or hypoplasia of the corpus callosum, and atrophy of the frontal and parietal cortex or arachnoidal cysts were reported. Patients with early-onset seizures had mild broadening of the ventricular system (P4), mild atrophy of frontal and parietal cortex (P7) and a retrocerebellar arachnoidal cyst or mega cisterna magna (P16). Additional less common clinical findings were scoliosis (26%), myopia (19%), strabismus (30%), hypogenitalism (19%) and accessory nipples (19%). Hands and feet were often described as slender and small, with single palmar creases in many and fetal pads in some patients. The lack of gross malformations is in accordance with the embryonic *tcf4* expression in D rerio, which is restricted to the brain and branchial arches.[250]

Most of the patients showed a happy and placid personality; violent or autoaggressive behavior was only reported in three patients. Stereotypic movements were observed in 30% of patients and in patient 13, the loss of hand use was reported. These features, in addition to microcephaly, breathing anomalies, severe mental retardation and seizures, resemble the features seen in both Rett and Angelman syndromes. Accordingly, most of the patients referred to us with suspected PTHS had already been tested for these diseases previously.

Table 1 Clinical findings in patients with *TCF4* mutations or deletions

Patient	Sex	Age	PTHS face	Severe MR	Micro-cephaly	Seizures	Hypo-tonia	breathing anomalies	MRI	Constip. HSCR	Palmar creases	Happy disp.	Scoliosis	Other anomalies
1	F	18y	+	+	+	-	+	-	-	+		+	+	FP, absent flexion crease thumb
2	M	17y	+	+	+	-	+	-	-	+		+		FP, small genitalia
3	M	2y	+	+	-	-	+	-	HCC	+	SPC	+	-	small genitalia
4	M	7y	+	+	-	Neonat.	+	+,4½ y	BV	+	SPC	+	-	ataxia, stereotypies
5	F	20y	+/-	+	-	-	+	+	-			outburst		
6	F	17y	+	+	+	-	+	+, 10y		+	SPC	+	+	locked left knee
7	F	20y	+	+	+	+, 9mo	+	+, first mo	FPA	+	SPC	+	+	myopia, strabismus
8	M	4y	+	+	+	+, 3y	+	+, 3y	-	+	SPC	+	-	FP, stereotypies
9	M	16y	+	+	+	+, 5y	-	+	ACC	+	SPC	+	-	FP, cryptorchidism
10	M	10y	+	+	-	++	-	+, 4y	-	-		+	+	AN
11	M	17y	+	+	-	-	-	+, 17y		+	SPC	+	-	pyloric stenosis, AN
12	M	2y 11mo	+	+	+	-	+	-	HCC, NC	+	SPC	+	-	frequent infections, myopia
13	M	14mo	+	+	-	-	+	-	-	-	simple	+	-	lost ability to use hands, stereotypies, FP of toes, cold hands and feet
14	F	18y	+	+	+	-	+	-	-	-	SPC	+	-	FP, ataxia, myopia, strabismus, stereotypies
15	F	2y 3mo	+	+	+	-	+	-	+	+	SPC	+ banging biting	-	myopia, hands in clenched posture, sleeping problems
16	F	2y 11mo	+	+	-	+, 4mo	+	+?	AC	-	SPC	+	-	recurrent ear infections
N=6[248]	4M 2F	8-29y	6/6	6/6	4/6	2/6	6/6	5/6 2-8y	4/4	3/6	4/5SPC 4/6SDFC	6/6 1 AB	2/6	1x AN, 2x strabismus
N=4[247]	2M 2F	4,5-10y	4/4	4/4	4/4	EEG 4/4		3/4 3,5-5y	3/3	4/4	+	4/4	1/4	4x stereotypies, 4x strabismus, 2x hypogenitalism, 2x AN
N=1[250]	F	6y	+	+	-	-	++	+, 7y	-	-	SPC	?	-	myopia, ataxia

AB, autoagressive behaviour; AC, arachnoidal cyst; AN, accessory nipple; BNC, bulging of caudate nuclei; BV, broadening of ventricles; constip., constipation; disp., disposition; EEG, electroencephalography anomalies; F, female; FP, fetal pads; FPA, frontal and pariental atrophy; HCC/ACC, hypoplasia/aplasia of corpus callosum; m, male; mo, months; NS, neonatal seizures; SDFC, supernumerary digital flexion creases; SPC, single palmar crease; y, years

As we found mutations in about 14% of the patients in this study, the important role of PTHS as a differential diagnosis of Rett and Angelman syndromes is further confirmed. The most consistent aspect distinguishing PTHS from the other two syndromes is the characteristic facial gestalt including a coarse face, high cheekbones, a beaked nasal tip, a protruding lower face, a wide mouth with cupid's-bow shaped upper lip and wide-spaced teeth. However, the facial phenotype can be subtle as that seen in patient 5, who does not have the characteristic beaked nose and resembles the other patients mostly through the shape of her face, with high cheekbones.

Mutational spectrum

In this study, we could identify 16 novel mutations in *TCF4*, distributed over the gene. (table 2, fig. 2); 13 of these are frameshift, nonsense or splice-site mutations, therefore further confirming TCF4 haploinsufficiency as the disease causing mechanism. One of these splice-site mutations is an A>G exchange at position IVS14+3, where both A and G are possible.[256] Nevertheless, because this mutation occurred *de novo*, we assumed abnormal splicing. We found an aberrant transcript by RT-PCR using cDNA from the patient, even though this was weaker than the wild-type allele, and confirmed skipping of exon 14 by sequencing the aberrant transcript (fig. 3B). One proven *de novo* frameshift mutation is located in the C-terminal exon 19 and results in an elongation of the putative protein by 37 amino acids. Owing to its location near to the C-terminus, nonsense-mediated mRNA decay, as assumed for stop mutations, might not occur, but impairment of protein function due to changes in protein structure is likely, particularly as the functional HLH domain is located in the C-terminal part of the protein.

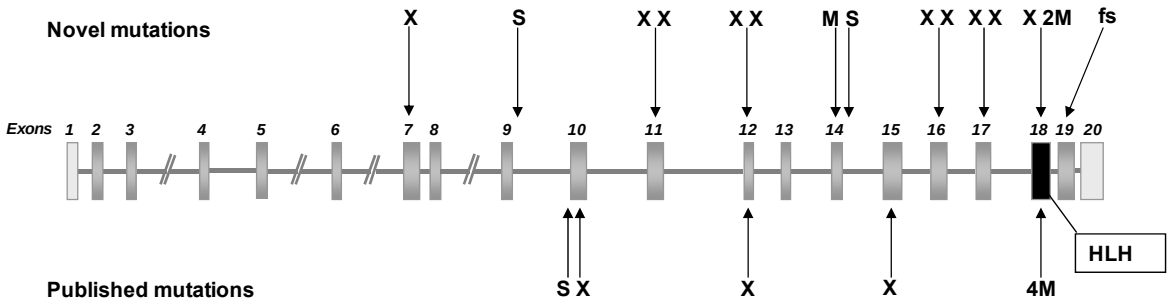

Figure 2 Scheme of the *TCF4* gene with localization of novel and previously published mutations.

Exons 2–19 contain coding sequence (dark-grey boxes), exons 1 and 20 are non-coding (light-grey boxes) and exon 18 encodes the helix–loop–helix domain (HLH, black box). fs, frameshift mutation leading to an elongated protein; M, missense mutation; S, splice-site mutation; X, stop mutation (nonsense or frameshift).

Amiel et al[247] had identified three missense mutations at the same amino acid position (576 or 580, depending on the isoform) within the bHLH domain in four patients, thus indicating a mutational hot spot at this position. Surprisingly, in 22 patients with defects in *TCF4* (7 and this study) we found this particular missense mutation only once (table 2). However, we found two further missense mutations in the bHLH domain located only two amino acid positions upstream of the reported recurrent mutation site, which lead to a change from arginine to proline or histidine at position 574 or 578, respectively. Both mutations at position 576/580 and 574/578 affect evolutionarily highly conserved glutamic and arginine residues constituting the E-box recognition motif.[247] We showed previously that such an impairment of the functional bHLH domain reduces interaction with ASCL1 in transactivating an E-box-containing reporter construct, to a similar degree as haploinsufficient stop mutations.[248]

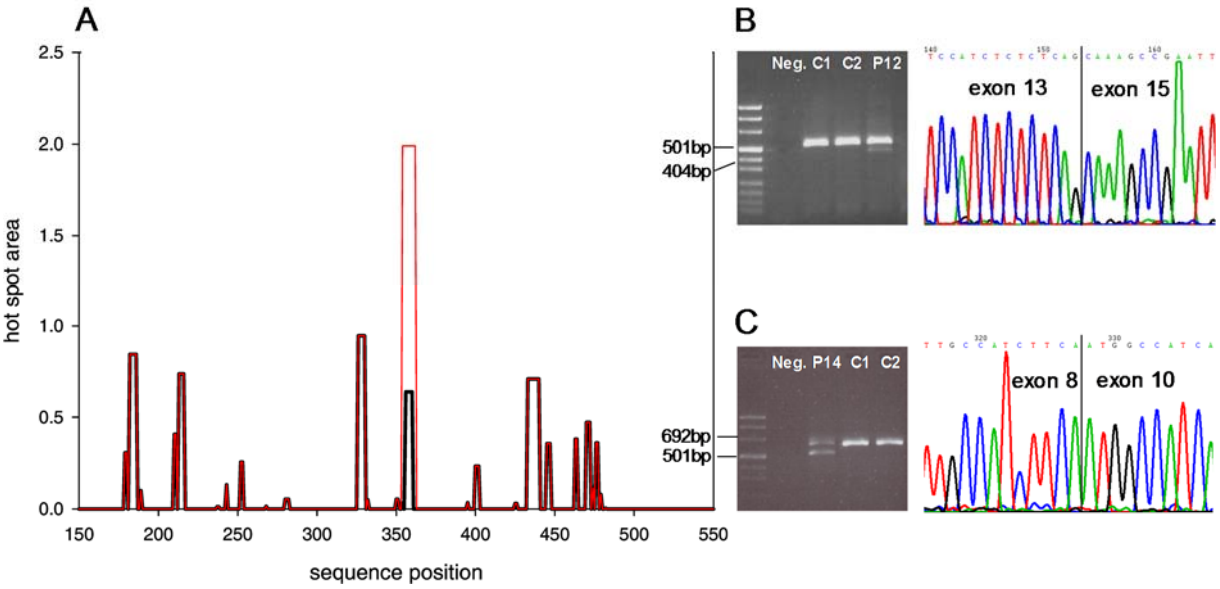

Figure 3 **A)** Calculated aggregation tendency of wild-type (black) and G358V (red) TCF4. The ‘‘hotspot area’’, which gives a measure of the aggregation tendency, was calculated using AGGRESCAN.[255] Compared with the size of adjacent hotspots, the hotspot emerging at the site of the G358V mutation is more than twice as large, suggesting an important effect of this amino acid exchange, probably leading to an increased formation of aggregates and therefore impaired function of TCF4. **B)** RT-PCR of patient 12 with the splice-site mutation IVS14+3ARG, showing the aberrant transcript in comparison with two normal control persons (C1, C2) and sequence of the aberrant transcript. **C)** Reverse transcriptase PCR of patient 14 with splice-site mutation IVS9+2insGT, showing the aberrant transcript in comparison to two normal control persons (C1, C2) and sequence of the aberrant transcript. Neg, negative control.

Interestingly, we also found the first missense mutation outside the bHLH domain. This G358V mutation in exon 14 affects an evolutionarily highly conserved position. The mutation was excluded in both parents and 192 healthy control chromosomes, and sampling errors

were excluded by short tandem repeat marker analysis. Nevertheless, as this mutation is not located within a known functional domain, the pathogenic mechanism remains unclear. Reporter-assay testing of interaction with ASCL1 showed a mild but not significant decrease in transcriptional activity (data not shown), which is not surprising as the mutation is not located within the DNA or protein-binding domain. Computational analyses indicated that the N-terminal 550 amino acids of TCF4 are predominantly disordered and do not adopt a globular (domain-like) tertiary structure. Such regions were reported previously to mediate

Table 2 Overview of novel and published *TCF4* defects

Exon	Genomic alteration	Protein alteration	Type	Frequency	Reference	This study
all	Deletion 1.2Mb		de novo	1	Zweier[248]	
all	Deletion 1.8Mb		de novo	1	Amiel[247]	
all	Deletion 0.5Mb		de novo	1	Brocksch.[250]	
7	c.469C>T	p.R157X	de novo	1		P5
IVS9	IVS9+2insGT	splice	de novo	1		P14
IVS9	IVS9-1G>C	splice	not tested	1	Zweier[248]	
10	c.692-694insT	p.G232fsX256	de novo	1	Zweier[248]	
11	c.791delG	p.S264fsX306	not tested	1		P11
11	c.908delC	p.T303fsX306	de novo	1		P2
12	c.937-940delG	p.314QfsX390	not tested	1		P15
12	c.949delA	p.S317fsX390	de novo	1		P6
12	c.965-969delATGCT	p.D322fsX336	de novo	1	Zweier[248]	
14	c.1073G>T	p.G358V	de novo	1		P7
IVS14	IVS14+3A>G	splice	de novo	1		P12
15	c.1153C>T	p.R385X	not tested	1	Zweier[248]	
16	c.1413-1414delG	p.V472fsX487	de novo	1		P4
16	c.1468-1471delC	p.Q491fsX534	de novo	1		P3
17	c.1512insA	p.S505fsX512	de novo	1		P10
17	c.1518delC	p.S507fsX534	de novo	1		P9
18	c.1687/1699A>T	p.K563/567X	not tested	1		P1
18	c.1721/1733G>C	p.R574/578P	de novo	1		P13
18	c.1721/1733G>A	p.R574/578H	ex. in moth	1		P16
18	c.1726/1738C>T	p.R576/580W	de novo	3	Zweier,[248] Amiel[247]	
18	c.1727/1739	p.R576/580Q	de novo	1	Amiel[247]	
19	c.1952-1957delCT	p.S661fs	de novo	1		P8

either specific interactions via short linear sequence motifs[254] or nonspecific interactions that might lead to aggregation.[255] However, computerized investigation of whether position 358 is part of a specific linear protein interaction motif gave no evidence for such a role of this sequence region (data not shown). In contrast, an analysis of the TCF4 aggregation properties revealed that the G358V mutation leads to an increase of the aggregation tendency, thus offering an explanation for the slightly reduced transcriptional activity observed. The small overall effect, however, suggests that only small-sized or transient/reversible aggregates are formed. This is also consistent with the immunofluorescence data that did not reveal visible TCF4 aggregates in cells transfected with the mutant protein compared with the wild type.

In conclusion, we have further delineated the mutational and clinical spectrum of PTHS and confirmed its important role in the differential diagnosis of severe mental retardation.

Acknowledgements

We are grateful to the patients and families who participated in this study. We thank D. Schweitzer for skillful technical assistance. This work was supported by grant RA 833/7-1 to A. Rauch funded by the Deutsche Forschungsgemeinschaft (DFG). M. Peippo at The Department of Medical Genetics, Väestöliitto is funded by Finland's Slot Machine Association (RAY).

Competing interests

None.

Patient consent

Parental consent obtained.

Chapter 4

CNTNAP2 and NRXN1 are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in Drosophila.

Zweier C, de Jong EK, Zweier M, Orrico A, Ousager LB, Collins AL, Bijlsma EK, Oortveld MA, Ekici AB, Reis A, Schenck A, Rauch A.

Am J Hum Genet. 2009 Nov;85(5):655-66.

Heterozygous copy-number variants and SNPs of *CNTNAP2* and *NRXN1*, two distantly related members of the neurexin superfamily, have been repeatedly associated with a wide spectrum of neuropsychiatric disorders, such as developmental language disorders, autism spectrum disorders, epilepsy, and schizophrenia. We now identified homozygous and compound-heterozygous deletions and mutations via molecular karyotyping and mutational screening in *CNTNAP2* and *NRXN1* in four patients with severe mental retardation (MR) and variable features, such as autistic behavior, epilepsy, and breathing anomalies, phenotypically overlapping with Pitt-Hopkins syndrome. With a frequency of at least 1% in our cohort of 179 patients, recessive defects in *CNTNAP2* appear to significantly contribute to severe MR. Whereas the established synaptic role of *NRXN1* suggests that synaptic defects contribute to the associated neuropsychiatric disorders and to severe MR as reported here, evidence for a synaptic role of the *CNTNAP2*-encoded protein CASPR2 has so far been lacking. Using *Drosophila* as a model, we now show that, as known for fly Nrx-I, the CASPR2 ortholog Nrx-IV might also localize to synapses. Overexpression of either protein can reorganize synaptic morphology and induce increased density of active zones, the synaptic domains of neurotransmitter release. Moreover, both Nrx-I and Nrx-IV determine the level of the presynaptic active-zone protein bruchpilot, indicating a possible common molecular mechanism in Nrx-I and Nrx-IV mutant conditions. We therefore propose that an analogous shared synaptic mechanism contributes to the similar clinical phenotypes resulting from defects in human *NRXN1* and *CNTNAP2*.

INTRODUCTION

The etiology of severe mental retardation (MR) is heterogeneous, and, despite a significant number of identified disease genes,[257] the majority of cases, especially non-syndromic cases, remain unsolved.[14] Many of the currently known MR-related genes are involved in neurogenesis and neuronal migration, and awareness of the implication of synaptic organization and plasticity in MR has only recently begun to rise.[4,258] In 2007, haploinsufficiency of the basic helix-loop-helix (bHLH) transcription factor 4 (TCF4) was identified as causative for Pitt-Hopkins syndrome (PTHS [MIM 610954]), a severe MR disorder with variable additional anomalies, such as breathing anomalies, epilepsy, and facial dysmorphism including a beaked nose and a wide mouth with a cupid's-bow shaped upper lip.[247,248] TCF4 belongs to the E-protein family of bHLH transcription factors, which bind as homo- and heterodimers to E-box consensus sequences in promoters of target genes.[259] Like other E-proteins, TCF4 shows a broad expression pattern and a high expression in the CNS.[250,260] After the identification of the underlying gene in 2007, approximately 50 patients have been reported,[247,248,250,260-262] demonstrating the importance of a diagnostic test for the

increased recognition and appreciation of a previously clinically underdiagnosed condition. Because of a similar severe degree of MR, commonly associated seizures, and microcephaly, PTHS has evolved as an important differential diagnosis to the two most common syndromic disorders in severe MR, Rett (MIM 312750) and Angelman (MIM 105830) syndromes.[262] Because only 12% of patients referred to us with suspected PTHS showed mutations in *TCF4* (Zweier et al.[262] and unpublished data), the clinically relatively homogenous group of 179 *TCF4*-mutation-negative patients, including two sibling pairs, represented a suitable study cohort for searching for additional candidate genes for overlapping disorders.

Through molecular karyotyping and mutational analysis, we indeed identified recessive defects in two genes, *CNTNAP2* and Neurexin I (*NRXN1*), in patients with a very similar severe MR disorder and variable additional symptoms, such as seizures and breathing anomalies, resembling Pitt-Hopkins syndrome. In light of the shared phenotype that characterizes our patients with recessive *CNTNAP2* and *NRXN1* defects, and on the basis of the theme of overlapping phenotypes being caused by genes that are linked with each other in molecular networks,[178] we further aimed to address the hypothesis of a common molecular pathogenesis. We therefore utilized the fruit fly *Drosophila melanogaster* as a model and collected data that point to a common synaptic link between these two genes.

SUBJECTS AND METHODS

Patients

Our study group consisted of 179 patients, including two sibling pairs, who were referred for *TCF4* testing because of severe MR and variable additional features reminiscent of the PTHS spectrum, such as microcephaly, dysmorphic facial gestalt, or breathing anomalies. *TCF4* mutational testing revealed normal results in all of these patients. Ethics approval for this study was obtained from the ethics committee of the Medical Faculty, University of Erlangen-Nuremberg, and informed consent was obtained from parents or guardians of the patients.

Molecular Karyotyping

Molecular karyotyping was performed in 48 patients with the Affymetrix 500 K SNPArray and in 12 patients with the Affymetrix 6.0 SNP Array, in accordance with the supplier's instructions. In the index patient of family 1, hybridization was performed with an Affymetrix GeneChip Mapping 500K SNP array, and the second affected patient and both parents were analyzed with the Affymetrix GeneChip Mapping 250K Nsp SNP array. Copy-number data

were analyzed with the Nexus software (Biodiscovery) and the Affymetrix Genotyping Console 3.0.2 software. Molecular karyotyping in patients 2 and 3 was performed with the Affymetrix GeneChip Mapping 6.0 array platform, and copy-number data were analyzed with the Affymetrix software Genotyping Console 3.0.2. The identified copy-number variants (CNVs) were submitted to the Decipher database (patient 1a, 250902; patient 2, 250903; patient 3, 250904).

Mutational screening

DNA samples from 177 patients, derived from peripheral-blood or lymphoblastoid cell lines, were screened for *CNTNAP2* (NM_014141) and *NRXN1* (NM004801) mutations by unidirectional direct sequencing of the coding exons 1–24 of *CNTNAP2* and the coding exons 2–22 of *NRXN1*, including intronic flanking regions (ABI BigDye Terminator Sequencing Kit v.3; Applied Biosystems), with the use of an automated capillary sequencer (ABI 3730; Applied Biosystems). Mutations were confirmed with an independent PCR and bidirectional sequencing. Primer pairs can be found in Table S1, available online. For splice-site prediction, the online tools NNSPLICE 0.9 and HSF V2.3 were used.

FISH and MLPA

Fluorescence in situ hybridization (FISH) analysis was performed in family 1 with the directly Cy3-labeled bacterial artificial chromosome (BAC) clone RP4-558L10 on metaphase spreads, in accordance with standard protocols.

Probes for all coding exons of *CNTNAP2* were designed and MLPA reaction was performed in accordance with the guidelines of MRC-Holland. The deletion in patient 3 was confirmed with MLPA with the use of a probe within exon 2 and a control probe within exon 12 of *NRXN1*. Probe sequences are listed in Table S2.

Analysis of relationship

The relationship of individuals within family 1 was analyzed, with a four-generation family with known relationships used as background, with the Graphical Representation of Relationships (GRR) software.[263] For GRR, we selected, from Affymetrix 250K arrays, 10,000 randomly distributed autosomal SNPs with a minimal minor allele frequency of 0.2 in Europeans. For each pair of individuals, GRR calculates over the 10,000 markers the identical by-state (IBS) mean and standard deviation. The graphical plot of IBS mean versus IBS standard deviation facilitates distinguishing between relationships such as parents and offspring, siblings, half siblings, and cousins, as well as identical or unrelated individuals. Additionally, SNP genotypes around *CNTNAP2* were analyzed in the family members.

Drosophila genes and lines

Drosophila orthologs of TCF4, NRXN1, and CASPR2 (daughterless [CG5102], Nrx-1 [CG7050], and Nrx-IV [CG6827]) were identified by the ENSEMBLE genome browser or by the reciprocal BLAST best-hit approach. Two RNAi lines, to Nrx-IV and daughterless, respectively, were obtained from the Vienna Drosophila Research Center (VDRC) and gave consistent phenotypes. VDRC lines no. 9039 (Nrx-IV) and no. 51297 (daughterless) were utilized for further analysis. RNA interference was induced with the UASGal4 system. The w1118 line (VDRC no. 60000) was used as a control, representing the same genetic background as the RNAi lines. Flies were raised at 28°C for maximum efficiency of knockdown. The Nrx-I overexpression line pUAST-Nrx-I was obtained from Wei Xie from Nanjing, China. Gal4 driver lines and the inducible Nrx-IV overexpression line P(EP)Nrx-IVEP604 were obtained from the Bloomington stock center.

Quantitative Real-Time PCR

RNA extraction from 333 L3 larvae of each genotype was performed with the RNeasy Lipid Tissue Kit (QIAGEN) in accordance with the supplier's protocols. cDNA synthesis was performed with iScript (Biorad). Quantitative real-time PCR was performed with the Power SYBR Green PCR Master Mix on a 7500 Fast Real-Time PCR System (Applied Biosystems), and results were normalized to the endogenous control actin. Primer sequences can be found in Table S3.

Immunostaining and data acquisition

We harvested 10–18 hr embryos and fixed them with 3.7% PFA for 20–25 min. All primary antibodies—anti-elav (labels nuclei of all neurons), antibody 22c10 (sensory nervous system), antibody BP102 (axon tracts, central neuropile region), anti-fas II (motor and central pioneer axons), and antibody nc82 (anti-bruchpilot, synaptic active zones) (all from the Developmental Studies Hybridoma Bank [DHSB])—were used in a 1:100 dilution. Late stage (16/17), nc-82-labeled embryos were assigned to one of three phenotypic categories: strong peripheral staining, moderate staining, and weak or residual staining, respectively. We performed statistical analysis of 69 wild-type (WT) embryos (w1118) and 73 Nrx-IV knockdown embryos from three independent experiments with a chi-square test and a Fisher's-exact test to obtain p values. The images for peripheral synaptic staining were obtained with a Zeiss Apotome.

Brains were dissected from L3 larvae and fixed for 30 min in 3.7% PFA. Pictures of WT and mutant brains were acquired with the use of the same microscope settings. Intensities of nc82 immunostainings were measured with Image J within two fields at two standardized positions in each CNS, one in the upper third and one in the lower third of the

ventral nerve cord. The average of these two values was normalized to the average of controls for comparison of results from independent experiments. A total of 45 w1118 brains from six independent experiments, 11 elav-Gal4::Nrx-I brains from two independent experiments, 19 double elav-Gal4::Nrx-I brains from three independent experiments, 12 elav-Gal4::Nrx-IV brains from two independent experiments, and 21 double-elav-Gal4::Nrx-IV brains from four independent experiments were measured. p values were obtained with a Wilcoxon test for two samples for comparison to the WT.

Type 1b neuromuscular junctions (NMJs) of muscle 4 were analyzed after dissection of L3 larvae and fixation in 3.7% PFA for 30 min. Costaining was performed with nc82 and DLG (both from DHSB) or HRP (Jackson Immuno Research) antibodies in a dilution of 1:500. NMJ pictures were stacked in ImageJ and processed in Adobe Photoshop. Numbers of active zones and branches were manually counted in an animated stack, and total synaptic area was determined by ImageJ. A total of 21 WT NMJs, 19 overexpression Nrx-I NMJs, and 20 overexpression Nrx-IV NMJs from at least two independent experiments were counted. For the evaluation of branches, 27 WT NMJs, 25 overexpression Nrx-I NMJs, and 20 overexpression Nrx-IV NMJs from three experiments were counted. We performed statistical evaluation with the Wilcoxon test for two samples, comparing each of the genotypes to the WT. The antibody against NrxIV was obtained from Christian Klämbt, Münster, Germany.[264] Secondary antibodies for all stainings were either Alexa 568- or Alexa 488-labeled antibodies against mouse or rabbit (Molecular Probes). All data were acquired blind to the evaluated phenotype.

RESULTS

Identification of recessive deletions and mutations in CNTNAP2 and NRXN1

Molecular karyotyping led to the identification of a homozygous deletion of exons 2–9 within the *CNTNAP2* gene on chromosome 7q35-q36.1 in a sibling pair of European origin (P1a and P1b), formerly published as possible clinical cases of Pitt-Hopkins syndrome.[228] This deletion was confirmed by FISH analysis (Figure S2) and MLPA (data not shown) and is predicted to be in frame but result in the loss of several functional domains (fig. 1A, fig. 2A, and fig. S1). Consanguinity of the parents had been denied,[228] and no indication for consanguinity was found by analysis of relationship with the use of the information of 10.000 SNPs. However, when the SNPs within and around *CNTNAP2* were analyzed, they showed homozygosity in both children, indicating an allele of common ancestry. By subsequent mutational screening of *CNTNAP2* in a larger cohort of 177 additional *TCF4*- mutation-negative patients, we identified a third patient of European origin (P2) with compound heterozygosity for the splice mutation IVS10-1G>T and a partial in-frame deletion of exons

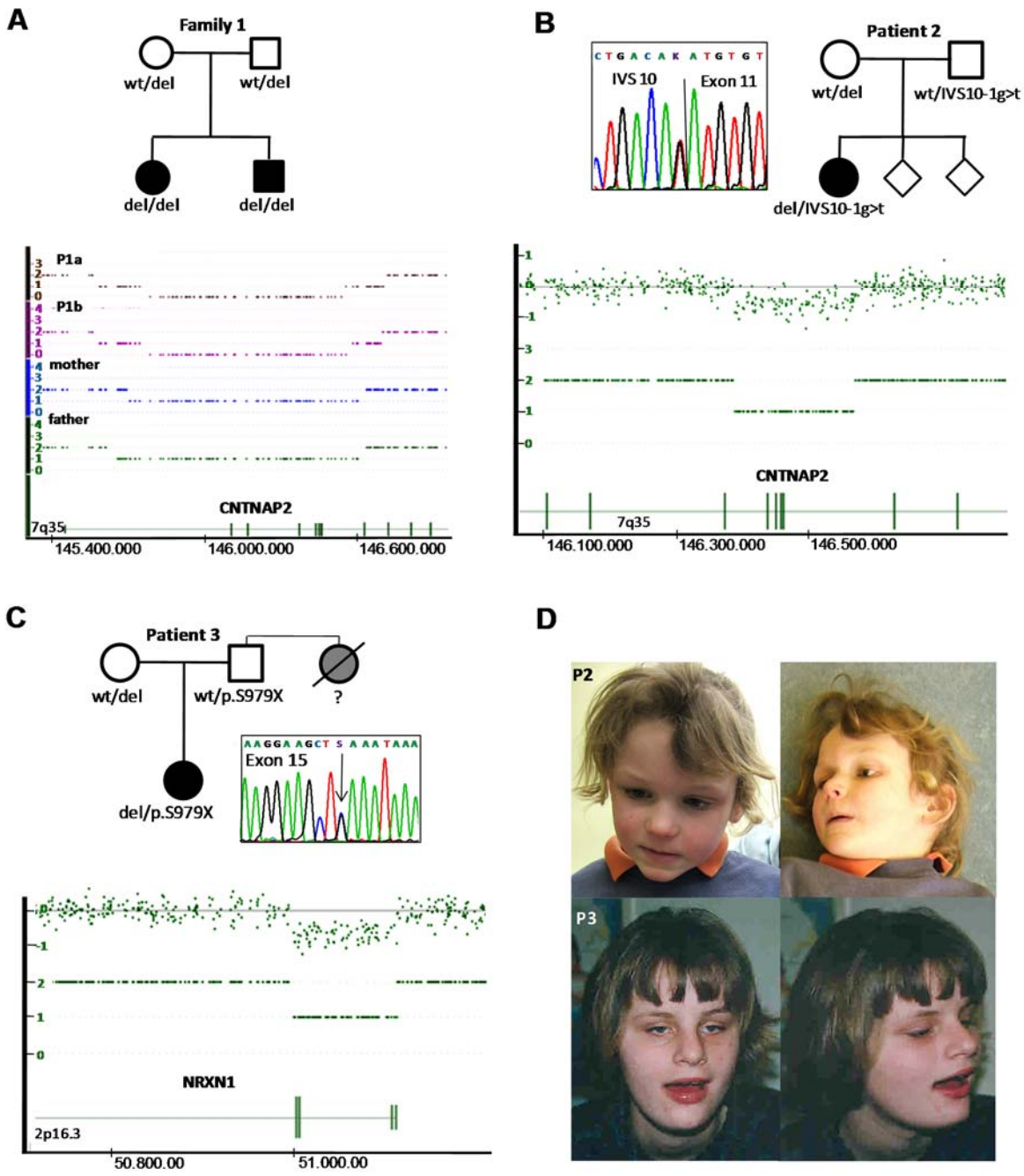

Figure 1 Pedigrees and results of molecular karyotyping.

A) Pedigree of family 1, with two affected children and homozygous deletion of *CNTNAP2* affecting exons 2–9. Both parents are heterozygous carriers of the deletion. Results are from molecular karyotyping with Affymetrix 250K SNP arrays and analysis with the Genotyping Console 3.0.2 software (Affymetrix). The deletion-flanking SNPs in the 500K array of P1a are SNP_A-1991616 (145,562,641 Mb; UCSC Human Genome Browser version 18 [hg18]) and SNP_A-1991672 (146,730,410 Mb; hg18), with a maximal deletion size of 1,167,269 bp and a minimal size of 1,146,016 bp (Nexus software). **B)** Pedigree of P2, with one affected child. The patient harbors an in-frame 180 kb deletion affecting *CNTNAP2* exons 5–8 and a splice-site mutation in the splice donor site of exon 11. Results of molecular karyotyping data from the Affymetrix 6.0 SNP array were analyzed with the

Genotyping Console 3.0.2 software, showing a deletion from CN_1217185 (146,387,354 Mb; hg18) to SNP_A-4269862 (146,566,863 Mb; hg18). See Figure S1 for SNP copy-number profiles.

5–8, identified by molecular karyotyping (fig. 1B and fig. 2A) and confirmed with MLPA. The splice-site mutation resulted in lack of recognition of the splice acceptor site by two splice-site-prediction programs and is therefore predicted to result in loss of exon 10, leading to a frameshift, and the deletion is predicted to result in the loss of two laminin G domains. The splice-site mutation was not found in 384 control chromosomes, and no *CNTNAP2* deletion was found in 667 molecularly karyotyped control individuals. In both families, the parents were heterozygous carriers of one of the respective defects.

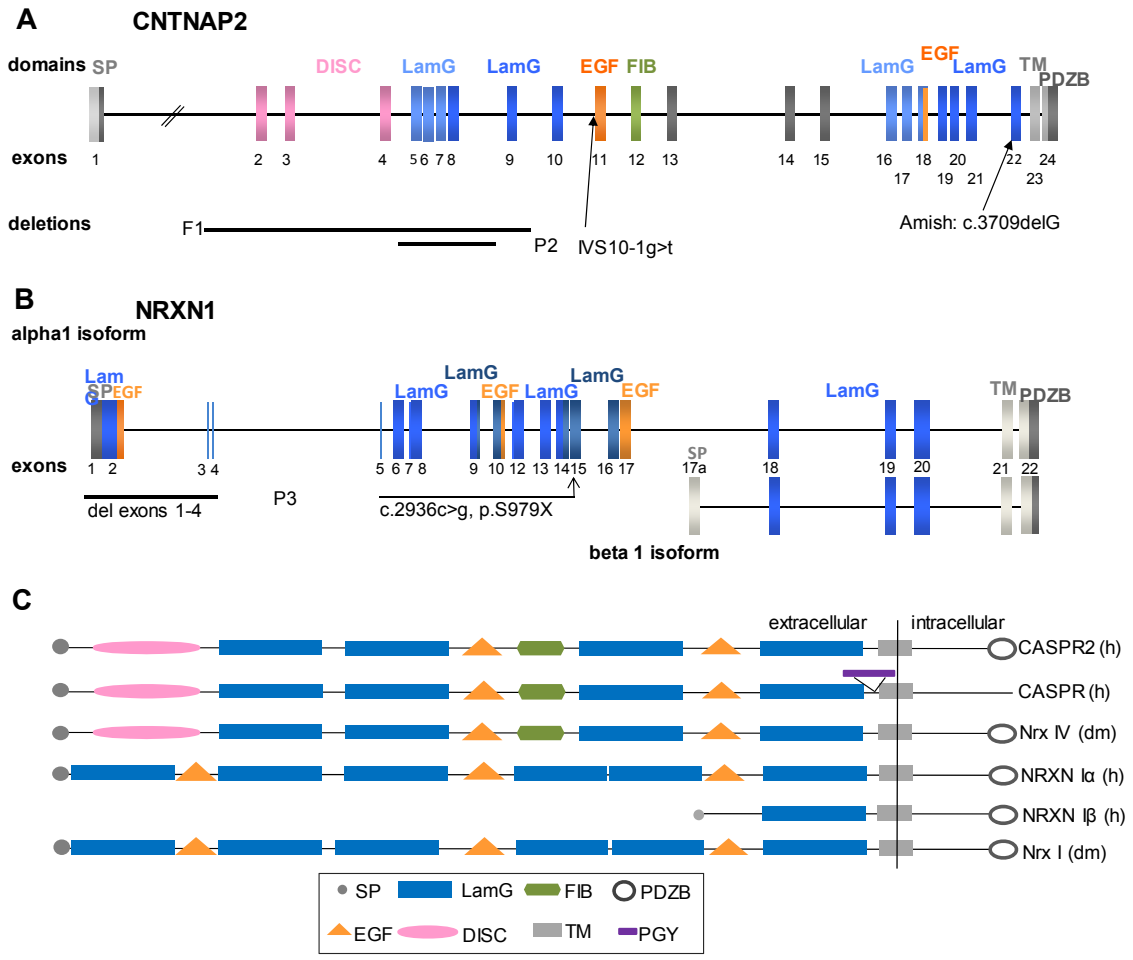

Figure 2 Structure of CNTNAP2 and NRXN1.

A) Schematic drawing of the genomic structure of *CNTNAP2* with color coding for domain-coding exons and localization of mutations and deletions. Black bars represent deletions. Abbreviations are as follows: SP, signal peptide; DISC, discoidin-like domain; LamG, laminin-G domain; EGF, epidermal growth factor-like domain; FIB, fibrinogen-like domain; TM, transmembrane region; PDZPB, PDZdomain-binding site ; F1, family 1; P2, patient 2; Amish, homozygous mutation in the Amish

population, published by Strauss et al.[184] **B)** Schematic drawing of the genomic structure of a-*NRXN1* and b-*NRXN1* with color coding for domain-coding exons and localization of the mutation and deletion in patient 3, the deletion being represented by a black bar. Abbreviations are as follows: SP, signal peptide; LamG, laminin-G domain; EGF, epidermal growth factor-like domain; TM, transmembrane region; PDZPB, PDZ-domain-binding site. **C)** Schematic drawing of the domain structure of neurexins, CASPR2, and CASPR in humans and *Drosophila*. In contrast to CASPR, CASPR2 contains a PDZ-domain-binding site but lacks a PGY repeat region, rich in proline, glycine, and tyrosine residues. Both neurexin I and CASPR2/Nrx-IV contain PDZ-domain-binding sites at their intracellular C terminus but differ in the presence of discoidin-like and fibrinogen-like domains and in the order of laminin-G domains.

In another European patient of our cohort who had a very similar phenotype (P3), we identified a heterozygous 180 kb deletion within the *NRXN1* gene on chromosome 2p16.3, spanning exons 1–4, including the start codon. This deletion was inherited from the healthy mother, but no deletions affecting the coding region of *NRXN1* were found in 667 molecularly karyotyped healthy controls. Subsequent sequencing of *NRXN1* in this patient revealed a stop mutation in exon 15 on the second allele, which was inherited from the healthy father (fig. 1C and fig. 2B). Both mutations are predicted to result in loss of the so-called alpha-isoform of NRXN1, one of two NRXN1 isoforms that are transcribed from alternative promoters. The presumably remaining shorter beta isoform (fig. 2B and 2C) appears not to be sufficient to ensure normal function, which is in accordance with findings in alpha-neurexin knockout mice.[265] Mutational screening of *NRXN1* in our study cohort did not reveal any additional defects.

Clinical characterization

As far as data are available, birth measurements of all patients (P1a, P1b, P2, and P3) were normal. Further growth development was also normal, apart from short stature in the siblings from family 1 and additional microcephaly in one of them. All four patients with recessive defects in *CNTNAP2* or *NRXN1* showed severe MR with lack of speech or with speech limited to single words (P1b), whereas motor milestones were normal or only mildly delayed, with a walking age of 2 years in P3. Episodes of hyperbreathing occurred in all patients, and seizures with an age at onset between 4 months and 30 months were observed in P1a, P1b, and P2. Additional variable anomalies were cerebellar hypoplasia, autistic behavior, and stereotypic movements. Apart from a wide mouth with thick lips in P1a and P1b and a wide mouth in P3, no specific facial dysmorphism were noted (Figure 1D). Parents of all patients were healthy, and the deceased sister of the father of P3 was said to have had epilepsy and mild MR. P1a and P1b have been described in detail by Orrico et al.,[228] and an overview of clinical details of all patients is shown in Table 1. Lack of *NRXN1* and *CNTNAP2* expression

in blood or fibroblasts (Bakkaloglu et al.[266] and data not shown) precluded functional studies on human material.

Table 1 Phenotype in patients with *CNTNAP2* and *NRXN1* mutations

	Siblings				
Patients	P1a	P1b	P2	Amish[184] (N=9)	P3
Mutations	*CNTNAP2* deletion exons 2-9, homozygous	*CNTNAP2* deletion exons 2-9, homozygous	*CNTNAP2* deletion exons 5-8 + IVS10-1G>T	*CNTNAP2* c.3709delG, homozygous	*NRXN1*deletion exons 1-4 + p.S979X
Age	20y	15y	11y	1-10y	18y
Sex	F	M	F	not reported	F
Parents	Healthy	healthy	healthy	not reported	healthy
Birth weight Length OFC	3700 g 51 cm 34,5 cm	not known	3700 g at term	not reported	3450 g normal normal
Height	<P3	<P3	normal	P4-P57	P50-75
Weight	P10	P5-10	P50	not reported	P50-75
OFC	<P3	P75	P75	P18-P99	P25
MR	Severe	severe	severe	all	severe
Age of walking	Normal	normal	not known	16-30m	2y
Speech	None	single words	none	yes, but regression	none
Seizures, age of onset	22mo	25-30mo	4-8mo	14-20mo	none
MRI	cerebellar hypoplasia	normal	not known	dysplasia in 43%	normal
Hyperbreathing	Yes	yes	yes	not reported	yes
Stereotypies	None	not noted	tooth grind., rep. hand movements		yes
Autistic behaviour	not noted	not noted	yes	67%	yes
Developmental regression	not noted	not noted	considered normal until 8 months	yes, with onset of seizures	normal the first years
Constipation	not noted	not noted	No	not reported	yes
Decreased deep tendon reflexes	not known	not known	not known	89%	UE: decreased LE: normal
Others	broad mouth, thick lips	broad mouth, thick lips	dry skin		broad mouth, strabismus, protruding tongue, excessive drooling , abnormal sleep-wake cycles, hypermotoric behavior
Normal testing[a]	*FRAXA, UBE3A, MECP2, TCF4*	*FRAXA, UBE3A, MECP2*	*UBE3A, CDKL5, MECP2, TCF4*		array CGH[b], *UBE3A, MECP2, ZEB2, TCF4*

Patients P1a and P1b are siblings from family 1, clinically described by Orrico et al.;[228] f, female; m, male; mo, months; y, years; MR, mental retardation; OFC, occipito-frontal-circumference; rep. hand movements, repetitive hand movements; tooth grind., tooth grinding; UE, upper extremities; LE, lower extremities; a, Previous genetic testing with reported normal results; b, Array-based comparative-genome hybridization via an Agilent 244K oligonucleotide array (Agilent)

Analysis of CNTNAP2 and NRXN1 orthologs Nrx-IV and Nrx-I in *Drosophila*

Although a synaptic role for *NRXN1* is known, this has not yet been established for *CNTNAP2*. However, the high similarity of clinical phenotypes caused by defects in the two genes suggested a potential common molecular contribution. To address this, as well as a further possible connection with TCF4, we utilized Drosophila as a model organism. All three genes, *TCF4*, *NRXN1*, and *CNTNAP2*, are highly conserved in evolution and have orthologs in Drosophila.

We initially hypothesized that the TCF4 ortholog daughterless might regulate Nrx-I and Nrx-IV as transcriptional targets. Knockdown of daughterless to 60% of WT levels by the use of two different ubiquitous driver lines (promoter-Gal4 lines that regulate inducible RNAi alleles; see Subjects and Methods) resulted in pupal lethality, confirming the importance of daughterless for fly development and viability.[267] However, expression of Nrx-I and Nrx-IV in L3 larvae of these genotypes was not significantly changed (fig. S3).

Drosophila Nrx-IV and Nrx-1 both determine the level of the synaptic protein bruchpilot

Knockdown of Nrx-IV, either ubiquitously with the use of actin-Gal4 drivers or specifically in neurons with the use of an elav-Gal4 driver, resulted in late embryonic lethality, with animals completing embryogenesis but failing to hatch. Together with an only recently reported expression of Nrx-IV in neurons,[264,268] this suggests a crucial role in this cell type. To evaluate the cause for lethality upon neuronal

knockdown, we performed immunostainings on embryos with a series of neuronal markers, labeling all neuronal nuclei, the sensory nervous system, main central axon tracts, and motor- and central pioneer axons, respectively. None of these showed abnormal position or morphology (fig. S4 and data not shown). However, we noted an overall diminished staining intensity of the presynaptic protein bruchpilot (nc82) in Nrx-IV knockdown embryos (fig. 3A). Quantification revealed that the fraction of embryos with weak or residual staining on peripheral motor synapses was significantly increased to 33% in Nrx-IV knockdown embryos, compared to 9% in WT embryos, whereas the fraction of strongly stained embryos was decreased to 25% in Nrx-IV knockdown embryos, compared to 55% in the WT (fig. 3B). Interestingly, bruchpilot is known to colocalize with Nrx-I at presynaptic active zones, the domains of neurotransmitter release,[269] and reduced levels of bruchpilot immunoreactivity have been reported in larval brains of Nrx-I mutants.[270] We studied bruchpilot levels after neuronal overexpression of either Nrx-I or Nrx-IV and found a significant dosage-dependent increase of bruchpilot-staining intensity in larval brains. Upon overexpression with one copy of a panneuronal elav-Gal4 driver, bruchpilot intensity was increased 1.2- and 1.3-fold, and

introduction of a second copy (dbElav) resulted in a 1.6- and 1.8-fold increase (fig. 3C and 3D).

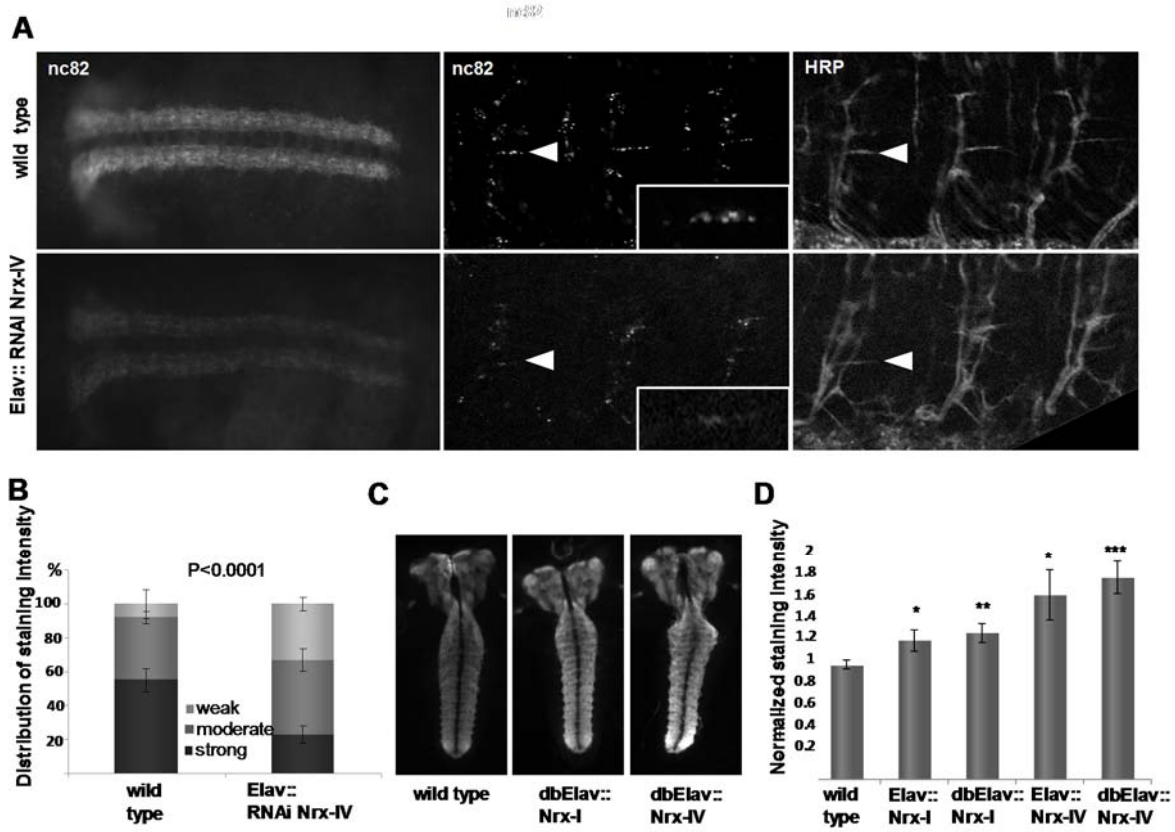

Figure 3 The presynaptic protein bruchpilot is misregulated in Nrx-IV knockdown embryos and larval brains with neuronal overexpression of Nrx-I and Nrx-IV.

A and **B)** For quantitative evaluation, embryos have been assigned to one of three categories of bruchpilot (nc82) intensities: strong, moderate, and weak peripheral staining. **A)** Images of strong and moderate central (ventral nerve cord, left panel) and peripheral (middle panel) presynaptic bruchpilot staining, representing the major fraction in WT and Nrx-IV knockdown embryos, respectively (UAS-Nrx-IVRNAi/þ; elav-Gal4/þ [Elav::RNAi Nrx-IV]). Note that peripheral and central synaptic staining of bruchpilot in the Nrx-IV knockdown embryos is diminished. Costaining with an anti-HRP marker, staining neuronal membranes, ensured the same focal plane in mutant and WT embryos. Arrowheads point to a specific identifiable synapse (the muscle 6/7 synapse). Bruchpilot staining of this synapse is depicted in insets in the middle panels. **B)** Quantitative analysis of nc82 (anti-bruchpilot, synaptic active zones) labeling. Diagram shows mean with standard deviation. **C)** Representative pictures of bruchpilot staining in brains with neuronal overexpression of Nrx-I or Nrx-IV as a result of the following genotypes: WT, UAS-Nrx-I/ elav-Gal4; elav-Gal4/þ (dbElav:: Nrx-I) and elav-Gal4/þ; UAS-Nrx-IV/ elav-Gal4 (dbElav:: Nrx-IV). Bruchpilot immunoreactivity is increased in both mutant conditions. **D)** Quantitative assessment of bruchpilot immunoreactivity in brains of genotypes shown in C). The diagram shows the mean of normalized intensity with the standard error of the mean. p values in (B) and (D) are related to the WT, respectively. Single asterisk, $p < 0.05$; double asterisk, $p < 0.01$; triple asterisk, $p < 0.0001$.

Drosophila Nrx-IV is present at synapses and can, like Nrx-I, reorganize them

For further analyses at the subsynaptic level, we utilized Drosophila larval NMJs, giant synapses that share a series of features with central excitatory synapses in the mammalian brain and represent an established model for the study of synaptic development and plasticity.[271] Staining of these synapses with a recently characterized specific antibody[264] for Nrx-IV detected the presence of Nrx-IV at synaptic terminals. Nrx-IV localizes in a pattern of subsynaptic foci that overlap active zones (fig. 4A), resembling the pattern previously reported for Nrx-I.20

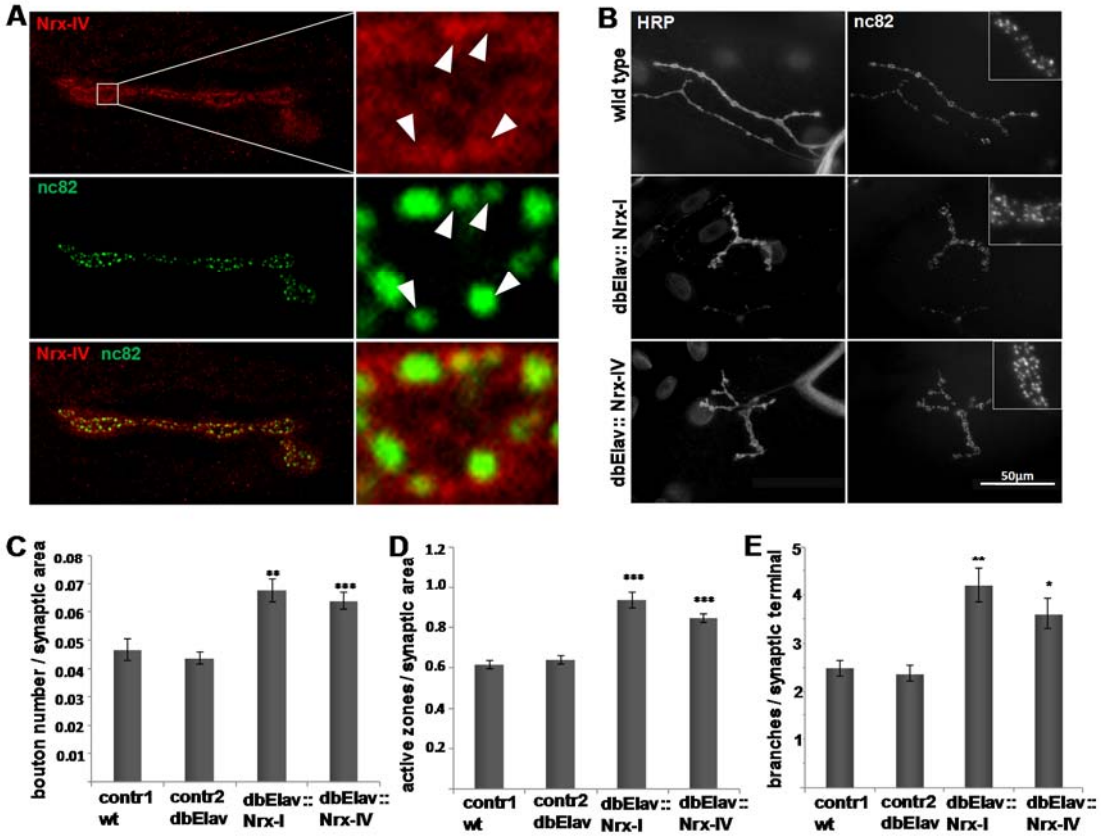

Figure 4 Nrx-IV localizes to synapses and its overexpression reorganizes synapse architecture. **A)** Presence of endogenous Nrx-IV at type 1b neuromuscular junctions (NMJs) of muscle 4, overlapping with active zones stained with nc82. White arrowheads point to Nrx-IV, labeling overlapping active zones. **B)** Synaptic terminal with increased bouton number and increased number of synaptic branches upon Nrx-I or Nrx-IV overexpression, stained with an anti-HRP marker and the nc82 antibody. The enlarged section depicts the increased density of active zones. **C)** Increased bouton number per mm2 NMJ area in UAS-Nrx-I/ elav-Gal4; elav-Gal4/þ (dbElav:: Nrx-I) and elav-Gal4/þ; UAS-Nrx-IV/elav-Gal4 (dbElav:: Nrx-IV) versus control w1118 animals (contr1) and dbElav control animals (contr2). **D)** Quantitative analysis of increased active-zone density (number per mm2 NMJ area) in neuronal Nrx-I and Nrx-IV overexpression. **E)** Increased number of synaptic branches in Nrx-I and Nrx-IV overexpression. Error bars indicate standard error of the mean. p values are related to the WT (contr1). Single asterisk, $p < 0.008$; double asterisk, $p < 0.0003$; triple asterisk, $p < 0.0001$.

Previous studies in Nrx-I null mutants revealed a decreased number of synaptic boutons in NMJs, whereas overexpression of Nrx-I resulted in an increased bouton number.[269]20 Our results, measuring total synaptic area in Nrx-I-overexpression animals, show that this goes along with a morphological reorganization into smaller boutons, as illustrated by an increased number of boutons per mm^2 NMJ area (fig. 4B and 4C). Strikingly, overexpression of Nrx-IV is capable of inducing the same morphological changes (fig. 4B and 4C). Furthermore, overexpression of either Nrx-I or Nrx-IV resulted in a highly significant increase in density of active zones (fig. 4B and 4D) and in a significant increase in branching of synaptic terminals (fig. 4B and 4E).

DISCUSSION

Recessive defects in CNTNAP2 and NRXN1 cause severe MR

We report here on homozygous and compound-heterozygous deletions and mutations in *NRXN1* and *CNTNAP2* that cause severe MR with additional features such as epilepsy, autistic behavior, and breathing anomalies. Heterozygous CNVs or SNPs in both genes have recently been extensively reported in association with autism-spectrum disorder (ASD) (AUTS15, [MIM 612100]), epilepsy, or schizophrenia.[266,272-282] Additionally, for *CNTNAP2*, an association with Gilles de la Tourette syndrome and obsessive compulsive disorder[283]4 was reported to be due to a disruption of the gene but was not confirmed in another family.[284] Furthermore, a homozygous stop mutation in *CNTNAP2* in Old Order Amish children was implicated in a distinct disorder, CDFE syndrome (MIM 610042), which is characterized by cortical dysplasia and early onset, intractable focal epilepsy leading to language regression, and behavioral and mental deterioration;[184] as well as periventricular leukomalacia and hepatomegaly in an additional patient.[285] Histological examination of temporal-lobe specimens of these patients showed evidence of abnormal neuronal migration and structure, as well as a possibly altered expression of Kv1.1 and Nav1.2 channels, therefore providing a possible explanation for the cortical dysplasia and epilepsy phenotypes.[184]

Taken together, published data and the present study indicate that heterozygous defects or variants in both *NRXN1* and *CNTNAP2* can represent susceptibility factors for variable cognitive, neurological, and psychiatric disorders, whereas biallelic defects result in a fully penetrant, severe neurodevelopmental disease such as that observed in our patients, thus representing different ends of the clinical spectrum caused by either monoallelic or biallelic defects in these two genes. This is in accordance with a report by Zahir et al.,[286] who described a patient with a heterozygous *de novo NRXN1* deletion affecting the same exons as those in our patient 3 but without detectable mutation on the second allele. This heterozygous deletion was associated with vertebral anomalies, behavioral problems, and only mild cognitive abnormalities. Of note, in our families, none of the heterozygous parents

had a history of autism, epilepsy, or schizophrenia, and none of other family members are known to have these disorders. However, the deceased sister of the father of patient 3 was said to have had epilepsy and mild MR. Whether or not this relates to the mutation in the father's family remains elusive.

A common feature in the Amish patients with a homozygous stop mutation and in our patients with homozygous deletions or compound-heterozygous deletion and mutation in *CNTNAP2* is the early onset of seizures. In contrast, none of our patients showed regression of speech development. Another difference is episodic hyperbreathing, present in our patients but not reported in the Amish patients. Cortical dysplasia, occurring in some of the Amish patients, was not observed in our patients; however, the number of patients examined with MRI is too low to allow any definite conclusions to be drawn. The phenotypic differences between our patients and the reported Amish patients may be explained by the loss of C-terminal transmembrane and cytoplasmic domains in the Amish mutation[184] and by the loss of N-terminal extracellular domains in our patients (fig. 2 and table 1). However, clinical bias cannot be excluded.

Resemblance among the patients in this study who have recessive defects in *CNTNAP2* or *NRXN1* is high, with the exception of patient 3, who has *NRXN1* defects but no seizures. It might even be speculated that the epilepsy, observed in patients with recessive *CNTNAP2* defects but not in our patient with the compound-heterozygous *NRXN1* defect, might be associated to the previously observed neuronal-migration anomalies,[184] whereas the remaining common symptoms such as severe MR and autistic behavior might be caused by overlapping synaptic anomalies. All of the reported patients were originally referred for *TCF4* analysis as a result of phenotypic resemblance to Pitt-Hopkins syndrome with regard to facial aspects, the severity of MR, and breathing anomalies. Nevertheless, phenotypical differences were noticeable. In contrast to patients with Pitt-Hopkins syndrome, who present an equally severe delay or lack of both motor and speech development, patients with *CNTNAP2* or *NRXN1* defects, also severely impaired in speech development, present with normal or only mildly to moderately delayed motor milestones. These findings are in line with the previously reported specific involvement of *CNTNAP2* in language development.[281]

Because of the similar phenotypes caused by defects in both *NRXN1* and *CNTNAP2* and the resemblance to Pitt-Hopkins syndrome, caused by haploinsufficiency of *TCF4*, we investigated an as-yet-unappreciated common molecular basis contributing to these disorders, using the fruitfly *Drosophila* as a model. The *Drosophila TCF4* ortholog daughterless belongs to the achaete-scute complex, which encodes members of the bHLH class of transcriptional factors with a well-established proneural function.[267] Our initial hypothesis, that daughterless might regulate Nrx-I and Nrx-IV as transcriptional targets, could not be confirmed by our analysis of their expression levels in the daughterless

knockdown condition (fig. S3). However, given that the data were obtained from knockdown animals, not null animals, this excludes neither transcriptional regulation that is undetectable by our assay nor other nontranscriptional interactions.

CNTNAP2 and NRXN1 *Drosophila* orthologs Nrx-IV and Nrx-I might be involved in a common synaptic mechanism

Our identification of a fully penetrant and severe MR phenotype due to recessive defects in NRXN1 is in agreement with NRXN1's important role in synaptic function, as previously suggested on the basis of its heterozygous alterations that are associated with neuropsychiatric disorders and work with animal models.[287] NRXN1 belongs to the evolutionarily conserved family of neurexins, presynaptic transmembrane proteins. Each of the three vertebrate neurexins has two promoters, generating longer alpha-neurexins and shorter beta-neurexins. In addition, extensive alternative splicing occurs, generating a large number of variants, which may mediate target recognition and synaptic specificity.[288] Alpha-neurexins also play a role in neurotransmitter release by coupling Ca2þ channels to synaptic-vesicle exocytosis.[289] In contrast to the three neurexins in mammals, there is only one Drosophila neurexin, termed Nrx-I or DNRX.[269,270] Previously, defects in several genes involved in synaptic pathways and complexes have been reported to be causative for or associated with MR or ASD, among them neurexin interactors. The extracellular region of neurexins binds to postsynaptic neuroligins, a family of proteins associated with autism and MR,[290] which in turn interact with the postsynaptic protein SHANK3, also associated with ASD.[291] The binding of neurexins to neuroligins connects pre- and postsynaptic neurons and mediates signaling across the synapse,[287] and the NRXN-NLGN-SHANK pathway is supposed to be crucial during synaptogenesis, given that mutations within this pathway lead to abnormal synaptogenesis and excess of inhibitory currents.[292] The intracellular domains of neurexins interact with the synaptic vesicle protein synaptotagmin and with PDZ-domain proteins such as CASK,[288] in which X-linked recessive mutations were only recently identified to be causative for MR and brain malformations.[133]

In contrast, a synaptic role of CNTNAP2 has not yet been established. CNTNAP2 encodes CASPR2, a protein related to neurexins. However, on the basis of its additional motifs, different domain organization,[288] and phylogenetic analyses, [269] only a distant relationship was assumed.[293] Caspr2 regulates neuron-glia contact in vertebrates and has been shown to colocalize with Shaker-like K+ channels in the juxtaparanodal areas of Ranvier nodes in myelinated axons of both the central and peripheral nervous system.[294,295] Until recently, its fly ortholog Nrx-IV was reported to be almost exclusively expressed in glial cells[296] and to regulate glia-glia contact in insects via septate junctions that are important for maintaining an intact blood-brain barrier.[293] A detected decrease in evoked neurotransmitter

release and an occasional failure to form synapses in Nrx-IV null mutants was interpreted as a consequence secondary to glial dysfunction.[296] However, recently, a neuronal Nrx-IV isoform was also identified.[264,268] Despite the knowledge of neuronal expression of vertebrate Caspr2,[297] so far only a report on detection of the protein in fractionated rat synaptic plasma membranes pointed to a possible synaptic presence.[266] Our findings now support a synaptic localization of Nrx-IV (Figure 4).

Previous studies at Drosophila neuromuscular junctions had reported a decreased number of synaptic boutons in Nrx-I null mutants and an increase upon overexpression of Nrx-I.[269] Our observation that the latter goes in hand with an increased active-zone density and the finding that Nrx-IV overexpression is causing the very same phenotypes (Figure 4) suggest a crucial and possibly common role for both proteins in the morphological organization of synapses. Moreover, published data for Nrx-I mutants[270] and our analysis of Nrx-IV knockdown and overexpression conditions of Nrx-I and Nrx-IV suggest an important role of both proteins in bidirectional regulation of bruchpilot levels. Bruchpilot is a presynaptic protein crucial for the structure of active zones, the subsynaptic domains of neurotransmitter release, and is present at most if not all synapses of Drosophila. It shows high sequence and functional homology to the vertebrate family of ELKS/CAST proteins,[298] corresponding to ERC1 and ERC2 in humans with similar reported functions. The mechanism by which Nrx-I and Nrx-IV determine bruchpilot levels remains the subject of future study. Given that all three contain C-terminal PDZ-domain binding sites, which are necessary for the formation of large complexes with active-zone proteins,[298] one reasonable hypothesis is that the three proteins might assemble into a synaptic complex or macromolecular network.

In conclusion, we have identified here autosomal-recessive defects in CNTNAP2 and NRXN1 in patients with severe MR and variable additional features overlapping with Pitt-Hopkins syndrome. With a frequency of at least 1% in this cohort, mutations in CNTNAP2 in particular appear to significantly contribute to severe MR. Using the fruitfly as a model organism, we observed that not only Nrx-I but also Nrx-IV, previously unrecognized as a synaptic protein in vertebrates and in Drosophila, is present at synapses and regulates levels of the active-zone protein bruchpilot. It is therefore tempting to hypothesize that misregulation of the human bruchpilot ortholog might underlie a common synaptic pathomechanism contributing to the similar phenotypes observed in patients with defects in CNTNAP2 and NRXN1.

Acknowledgments

We thank the patients and families for participation in this study and Christine Zeck-Papp, Michaela Kirsch, and Daniela Schweitzer for excellent technical assistance. We are grateful to Wei Xie (Genetics Research Center, Southeast University Medical School, Nanjing, China)

and to Christian Klämbt (Institute for Neurobiology, Münster, Germany) for providing the pUAST-Nrx-I flies and the Nrx-IV antibody, respectively. We thank H. Brunner and H. van Bokhoven for critical reading of the manuscript. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) to C.Z.; by a grant from the German Mental Retardation Network (MRNET) funded by the German Federal Ministry of Education and Research (BMBF) as a part of the National Genome Research Network (NGFNplus), to A.Re., A.Ra., and A.S.; and by a VIDI grant from the Netherlands Organization for Scientific Research (NWO) to A.S.

Web Resources

The URLs for data presented herein are as follows:

Basic Local Alignment Search Tool (BLAST), http://blast.ncbi.nlm.nih.gov/Blast.cgi

Database of Drosophila Genes and Genomes, http://flybase.org

Decipher database, http://decipher.sanger.ac.uk/application/

Graphic Representation of Relationships,
http://www.sph.umich.edu/csg/abecasis/GRR/index.html

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

Splice site prediction, http://www.fruitfly.org/seq_tools/splice.html and
http://www.umd.be/HSF/

University of California, Santa Clara (UCSC) Genome Browser, http://genome.ucsc.edu

SUPPLEMENTAL DATA

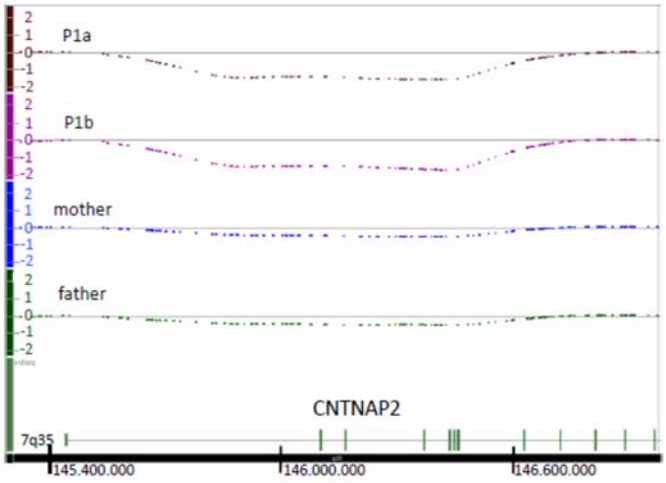

Figure S1 SNP copy number profile in family 1

SNP copy number profile (Log2 ratio) of the Affymetrix 250 K array, analyzed with the Genotyping Console 3.0.2 software (Affymetrix) in all members of family 1, showing a homozygous deletion within the *CNTNAP2* gene in both affected children and a heterozygous deletion in both healthy parents.

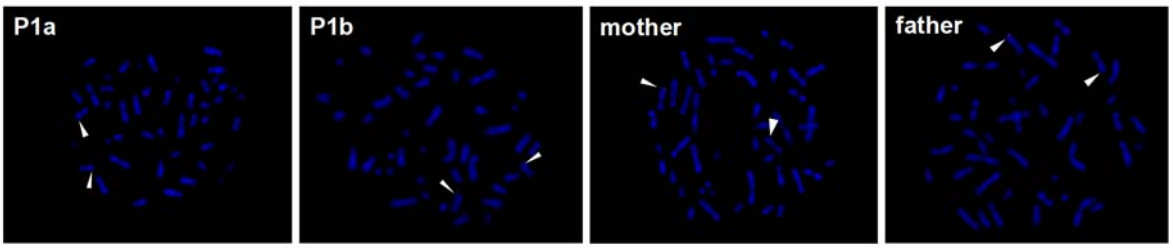

Figure S2 FISH analysis in family 1

FISH analysis with the directly Cy3-labeled BAC clone RP4-558L10 on metaphase spreads. White arrowheads indicate chromosomes 7. In the patients (P1a, P1b) a homozygous deletion of the BAC clone can be seen, in both parents a heterozygous deletion each.

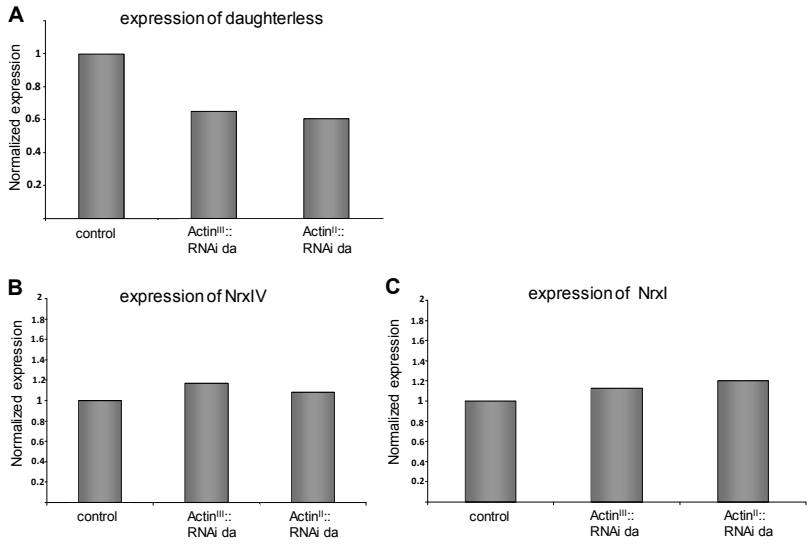

Figure S3. Expression of daughterless, Nrx-IV and Nrx-I in daughterless knockdown larvae

A) Normalized expression of daughterless after a knockdown with two different ubiquitous drivers on the second (II, Actin-Gal4 17bFO1) and third (III, Actin-Gal4 25FO1) chromosome, respectively. In both cases the knockdown is down to 60% compared to the wt (w1118) **B,C)** Normalized expression of Nrx-IV and Nrx-I upon knockdown of daughterless, respectively. No significant changes were detected.

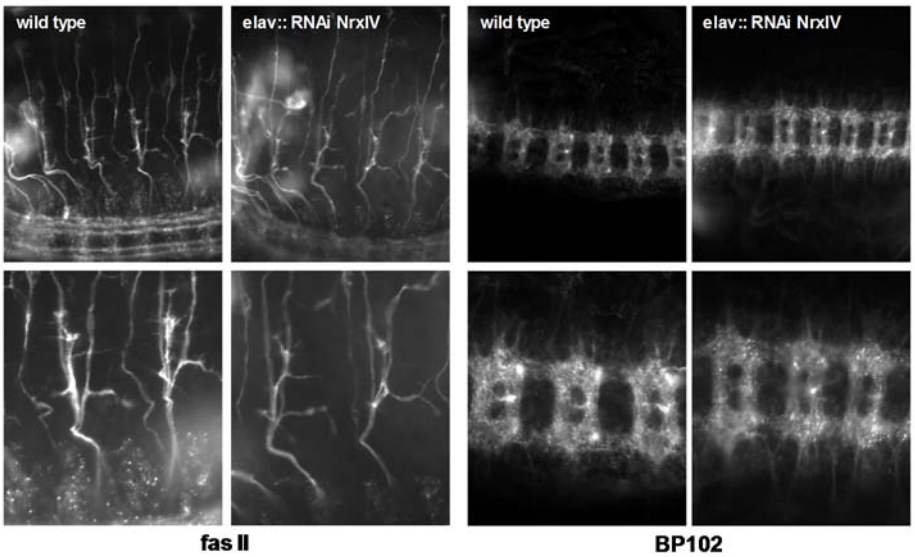

Figure S4 Immunostaining of the embryonic nervous system

Examples of immunofluorescence analysis of the nervous system in wild type and Nrx-IV panneuronal knockdown embryos. Anti-fas II stains motor- and central pioneer axons. Antibody BP102 stains axon tracts/the central neuropile region. No apparent morphological or structural alterations were observed in the Nrx-IV knockdown embryos compared to the wild type.

Table S1 Primer sequences for amplification and sequencing of all coding exons of CNTNAP2 and NRXN1

Exon	primer sequence	amplicon length (bp)
CNTNAP2_e1F	CCTCGCGTATTTGAGGACAG	345
CNTNAP2_e1R	AGTGGCTGCAAGTGTGTGAC	
CNTNAP2_e2F	GAATTGCCTAAATTCCTTTGC	363
CNTNAP2_e2R	TGGTGTCTGCCAACATCTG	
CNTNAP2_e3F	GCACTGCCAAGACCAATTAAG	427
CNTNAP2_e3R	TGATGAATAAATAGTTTCCCAATG	
CNTNAP2_e4F	Catggatgaaaagacccaca	481
CNTNAP2_e4R	Aaggtagtttattgtcagagaaagca	
CNTNAP2_e5F	Tctttgcagacacctgttgg	452
CNTNAP2_e5R	Tttttgaatgactaggtgtttcatt	
CNTNAP2_e6F	Tcccaggttaactcgaatgg	477
CNTNAP2_e6R	Tgaaacgaattaatcaggttttt	
CNTNAP2_e7F	GCCATAGATTTTGGAGGCAG	413
CNTNAP2_e7R	ACATCATTTTGCCCAAACAC	
CNTNAP2_e8F	TCACTGAATCCATGCTCTGC	524
CNTNAP2_e8R	AAAACCTAATCCTGAGCGTGTAAC	
CNTNAP2_e9F	Ttgtaagcagcactgtattttcc	497

CNTNAP2_e9R	Ggccagaagaatatggtgaca	
CNTNAP2_e10F	GAAACAGTAGTTGGATGTGATGG	408
CNTNAP2_e10R	GAATGGTAATTTCCACCTTACCTG	
CNTNAP2_e11F	CCTTGGTAAGGCAACCTGG	365
CNTNAP2_e11R	GAAATGACAATTGGAATCTTGG	
CNTNAP2_e12F	CTCTTTCCAGGAAGAACTACTCC	369
CNTNAP2_e12R	GCAATATGTTGCTGATTAGATGTTG	
CNTNAP2_e13F	GCTCTCCTTAACACTGTTCTACACC	460
CNTNAP2_e13R	CTTATTTACAGCTTCCTTCCTACTG	
CNTNAP2_e14F	Agagtattcctggggaagtgg	440
CNTNAP2_e14R	Ttgtcgcactgacctctttct	
CNTNAP2_e15F	CCAAACGATTACTGAAATGTCATC	373
CNTNAP2_e15R	ATCTCTGCTTGGGTTGTGTG	
CNTNAP2_e16F	Tgtgaggatttggtccaatg	469
CNTNAP2_e16R	Aggcttgtgtgtccacctct	
CNTNAP2_e17F	TCGACCTTTGTAGGACGTGAC	479
CNTNAP2_e17R	GGCCAACACCTTTACTTTTGG	
CNTNAP2_e18F	GCCTTATAGCCTGCAGGGAG	587
CNTNAP2_e18R	GATTAGGAAATGATTTTGGTTGG	
CNTNAP2_e19F	TACTCAGATGCCCTTCCTGG	462
CNTNAP2_e19R	GCCTATGGGGAATAAATAACAAAC	
CNTNAP2_e20F	Agcaggaattgaggggatgt	350
CNTNAP2_e20R	Ttatgcacttgtaggagaaagtgt	
CNTNAP2_e21F	GAAAACCAGGGTTCAAAGAGTG	314
CNTNAP2_e21R	AAGATATTCGTGACTGGCCC	
CNTNAP2_e22F	Gctttggacacaagcattca	462
CNTNAP2_e22R	Acgttcctttgccctttctt	
CNTNAP2_e23F	Gttgtgattcttgtgggagaca	366
CNTNAP2_e23R	Cagcaaaatgaataatgtaaaaacc	
CNTNAP2_e24F	GAGAGGGCTGTGTCTGACG	437
CNTNAP2_e24R	ATATTCCATTGCCTGCCTCC	
NRXN1_ex2aF	CCCTTACCTTTCTGTCTCTCG	499
NRXN1_ex2aR	GTCGATGAAGAGCGTGGTGT	
NRXN1_ex2bF	CGACTTCCTGGAGCTGATTC	696
NRXN1_ex2bR	Gagtcccccagaaacaaggt	
NRXN1_ex3F	Tgattttgttttccccttg	344
NRXN1_ex3R	Gatggcagggttgagaatgt	
NRXN1_ex4F	Gtgcagcatatgccagtgtt	446
NRXN1_ex4R	Agccacaggaacaaaccaaa	
NRXN1_ex5F	Aagccaggctgtctctgcta	386
NRXN1_ex5R	Tctggattggtctcgaggtt	
NRXN1_ex6F	Ctgtttgactgagacagagttca	596
NRXN1_ex6R	Tggcaggaagattcatctcag	
NRXN1_ex7F	Ctctgtgggaggcctacttg	418
NRXN1_ex7R	Cagatgaaaagaaggaggtcaaa	
NRXN1_ex8F	Caggcatatcccaggattaca	341
NRXN1_ex8R	Gtgccgtttgactctggaac	
NRXN1_ex9F	Tcgttgaaagttacatgagctg	595
NRXN1_ex9R	Tcacttttaggaatggcatgg	
NRXN1_ex10F	Ttgtctgcctccaaggagtt	573
NRXN1_ex10R	Caatggtcagtgcaggtttg	
NRXN1_ex11F	Tgaagaagatgaattgattttgtagt	493
NRXN1_ex11R	Cccccgaaaacctcaaatta	
NRXN1_ex12F	Ttggccatttttaaaaccttc	432
NRXN1_ex12R	Gcagtgggaaagtcttcagc	
NRXN1_ex13F	Ctctgtgaggttcatttgcttg	380
NRXN1_ex13R	Tgtcatttttgaggaaaaacacc	
NRXN1_ex14F	Tgtgatattgttatgaaagcctaaa	599
NRXN1_ex14R	Tgccttcttatttgtcctcca	
NRXN1_ex15F	Ggatggaaccacctgaaaaa	450
NRXN1_ex15R	Cagaggcttgtgtgttgtattg	
NRXN1_ex16F	Tttagcactttggggaaaaca	433
NRXN1_ex16R	Ccaaattgggtatttgaccag	

NRXN1_ex17F	Cagcctctcagttcctaattca	429
NRXN1_ex17R	Cctttccgtagaaacaactgc	
NRXN1_ex18F	Gccatagttttgtgtgtagagttga	416
NRXN1_ex18R	Tttcctataacaaagtactggtttctg	
NRXN1_ex19F	Gaagtaaaattggaggaaagca	450
NRXN1_ex19R	Ggacagcattacattcacatgac	
NRXN1_ex20F	Tgctctcattattcaccccata	541
NRXN1_ex20R	Ggaagctgtagtgcctaagatca	
NRXN1_ex21F	Aaagggaaatagtgaatttggtttc	449
NRXN1_ex21R	Aagccctgtgtgctataccc	
NRXN1_ex22F	Aggcaaagggatggctacat	550
NRXN1_ex22R	TGTGCTTCATAAAAAGGAAAGTAAA	

Table S2 Probe sequences for MLPA analysis of all coding exons of *CNTNAP2* and exon 2 of *NRXN1*

exon	probe sequence	probe length
CNTNAP2_ex1_F	GGGTTCCCTAAGGGTTGGATGCTGTGGATTGTCAGCAGCTGCCTCT	96
CNTNAP2_ex1_R	GCAGAGCCTGGACGGCTCCCTCCACGTTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex2_F	GGGTTCCCTAAGGGTTGGAGTGATGAGCCACTTGTCTCTGGACTCCCCCA	104
CNTNAP2_ex2_R	TGTGGCTTTCAGCAGCTCCTCCTCCATCTCTTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex3_F	GGGTTCCCTAAGGGTTGGAGTCTCCATCAGACAGCGACCATTATCAATGGCT	108
CNTNAP2_ex3_R	TCAGGTTGACTTTGGCAATCGGAAGCAGATCAGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex4_F	GGGTTCCCTAAGGGTTGGAGAAACATTAACTCTGACGGTGTGGTCCGGCACGAA	112
CNTNAP2_ex4_R	TTACAGCATCCGATTATTGCCCGCTATGTGCGCATTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex5_F	GGGTTCCCTAAGGGTTGGAGATGGCCATGTTGTATTACCATATAGATTCAGAAACAA	116
CNTNAP2_ex5_R	GAAGATGAAAACACTGAAAGATGTCATTGCCTTGAATCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex6_F	GGGTTCCCTAAGGGTTGGAGCATTAACCTCACTCTGGACAGGAGCATGCAGCACTTCC	120
CNTNAP2_ex6_R	GTACCAATGGAGAGTTTGACTACCTGGACTTGGACTATGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex7_F	GGGTTCCCTAAGGGTTGGAGAGGCATCCCTTTCTCTGGCAAGCCCAGCTC	104
CNTNAP2_ex7_R	CAGCAGTAGAAAGAATTTCAAAGGCTGCATGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex8_F	GGGTTCCCTAAGGGTTGGAGGACGGCTTAACCAGGACCTGTTCTCAGTCAGTTTCCAGTT	124
CNTNAP2_ex8_R	TAGGACATGGAACCCCAATGGTCTCCTGGTCTTCAGTCACTTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex9_F	GGGTTCCCTAAGGGTTGGAGAATGATGGACAGTGGCACGAGGTTCGCTTCCT	108
CNTNAP2_ex9_R	AGCCAAGGAAAATTTTGCTATTCTCACCATCGATCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex10_F	GGGTTCCCTAAGGGTTGGACATTCCAAGGATGCATGCAGCTCATTCAAGTGGA	112
CNTNAP2_ex10_R	CGATCAACTTGTAAATTTATACGAAGTGGCACAAAGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex11_F	GGGTTCCCTAAGGGTTGGAGACAGCTTCAAATGCACTTGTGATGAGAC	100
CNTNAP2_ex11_R	AGGATACAGTGGGGCCACCTGCCACAACTTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex12_F	GGGTTCCCTAAGGGTTGGATATCTACGAGCCTTCCTGTGAAGCCTACAAA	104

CNTNAP2_ex12_R	CACCTAGGACAGACATCAAATTATTACTGGATCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex13_F	GGGTTCCCTAAGGGTTGGAGCAGATGCAGACGCCTGTGGTCGGCTACAACCCAGAA	116
CNTNAP2_ex13_R	AAATACTCAGTGACACAGCTCGTTTACAGCGCCTCCATCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex14_F	GGGTTCCCTAAGGGTTGGAGCCCTTACACTTGGTGGGTTGGCAAAGCCAACG	108
CNTNAP2_ex14_R	AGAAGCACTACTACTGGGGAGGCTCTGGGCCTGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex15_F	GGGTTCCCTAAGGGTTGGATTGGAGATACTGACCGTCAAGGCTCAG	96
CNTNAP2_ex15_R	AAGCCAAATTGAGCGTAGGTCCTCTGCTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex16_F	GGGTTCCCTAAGGGTTGGAGAAACTAGCGCTGACATTTCTTTCTACTTCAAAAC	112
CNTNAP2_ex16_R	ATTAACCCCCTGGGGAGTGTTTCTTGAAAATATGGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex17_F	GGGTTCCCTAAGGGTTGGAGTGTCCTTTTCATTTGATGTGGGAAATGGGCCAGTAGAG	120
CNTNAP2_ex17_R	ATTGTAGTGAGGTCACCAACCCCTCTCAACGATGACCAGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex18_F	GGGTTCCCTAAGGGTTGGACAAAGGTCACATCTGGGTTCATATCCGGATGCTCGGGCCAT	124
CNTNAP2_ex18_R	TGCACCAGCTATGGAACAAACTGTGAAAATGGAGGCAAATGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex19_F	GGGTTCCCTAAGGGTTGGACCTGGCACAGGAGGAGATCCGCTTCAGCTTCAGCACCACCA	124
CNTNAP2_ex19_R	AGGCGCCCTGCATTCTCCTCTACATCAGCTCCTTCACCACATCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex20_F	GGGTTCCCTAAGGGTTGGATGGGTGGCACCCGAGAGCCATACAATATT	100
CNTNAP2_ex20_R	GACGTAGACCACAGGAACATGGCCAATGGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex21_F	GGGTTCCCTAAGGGTTGGATCCTTCTGTGAGTTACCATCTGCCAAGTT	100
CNTNAP2_ex21_R	CATCCGACACCCTCTTCAATTCTCCCAAGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex22_F	GGGTTCCCTAAGGGTTGGACAGGATTCACTGGTTGCCTCTCCAGAGTCCAGTTCAACC	120
CNTNAP2_ex22_R	AGATCGCCCCTCTCAAGGCCGCCTTGAGGCAGACAAACGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex23_F	GGGTTCCCTAAGGGTTGGAGACAAGGCCAAGCTATAAGAAATGGAG	96
CNTNAP2_ex23_R	TCAACAGAAACTCGGCTATCATTGGAGTCTAGATTGGATCTTGCTGGCAC	
CNTNAP2_ex24_F	GGGTTCCCTAAGGGTTGGACTGTGGTGATTTTCACCATCCTGTGCACCCTGGTCTT	116
CNTNAP2_ex24_R	CCTGATCCGGTACATGTTCCGCCACAAGGGCACCTACTCTAGATTGGATCTTGCTGGCAC	
NRXN1_ex2_F	GGGTTCCCTAAGGGTTGGACGTGAGGGTCAACTCCTCGCAGG	88
NRXN1_ex2_R	TCCTGCCCGTGGACAGCGGCGAGTCTAGATTGGATCTTGCTGGCAC	

Table S3 Primer sequences for quantitative real-time PCR

Gene	primer sequence	product length (bp)
Daughterless_F	CTCGCTGCAACAAAAGGAAT	120
Daughterless_R	AAGCAGTTCTGGAACACCTCA	
Neurexin-I_F	AATCTGCGGCTGCAAGTC	105
Neurexin-I_R	GAAGAGACCACCCAGGTGAA	
Neurexin-IV_F	TGCCATACATCAAACAATCCA	111
Neurexin-IV_F	AGGTTCTAGGGGACCACTGC	

Chapter 5

Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1.

Gregor A, Albrecht B, Bader I, Bijlsma EK, Ekici AB, Engels H, Hackmann K, Horn D, Hoyer J, Klapecki J, Kohlhase J, Maystadt I, Nagl S, Prott E, Tinschert S, Ullmann R, Wohlleber E, Woods G, Reis A, Rauch A, Zweier C.

BMC Med Genet. 2011 Aug 9;12:106.

ABSTRACT

Background

Heterozygous copy-number and missense variants in *CNTNAP2* and *NRXN1* have repeatedly been associated with a wide spectrum of neuropsychiatric disorders such as developmental language and autism spectrum disorders, epilepsy and schizophrenia. Recently, homozygous or compound heterozygous defects in either gene were reported as causative for severe intellectual disability.

Methods

99 patients with severe intellectual disability and resemblance to Pitt-Hopkins syndrome and/or suspected recessive inheritance were screened for mutations in *CNTNAP2* and *NRXN1*. Molecular karyotyping was performed in 45 patients. In 8 further patients with variable intellectual disability and heterozygous deletions in either *CNTNAP2* or *NRXN1*, the remaining allele was sequenced.

Results

By molecular karyotyping and mutational screening of *CNTNAP2* and *NRXN1* in a group of severely intellectually disabled patients we identified a heterozygous deletion in *NRXN1* in one patient and heterozygous splice-site, frameshift and stop mutations in *CNTNAP2* in four patients, respectively. Neither in these patients nor in eight further patients with heterozygous deletions within *NRXN1* or *CNTNAP2* we could identify a defect on the second allele. One deletion in *NRXN1* and one deletion in *CNTNAP2* occurred *de novo*, in another family the deletion was also identified in the mother who had learning difficulties, and in all other tested families one parent was shown to be healthy carrier of the respective deletion or mutation.

Conclusions

We report on patients with heterozygous defects in *CNTNAP2* or *NRXN1* associated with severe intellectual disability, which has only been reported for recessive defects before. These results expand the spectrum of phenotypic severity in patients with heterozygous defects in either gene. The large variability between severely affected patients and mildly affected or asymptomatic carrier parents might suggest the presence of a second hit, not necessarily located in the same gene.

BACKGROUND

Recent data suggested that heterozygous variants or defects in *NRXN1* (Neurexin 1) or *CNTNAP2* (contactin associated protein 2), both genes encoding neuronal cell adhesion molecules, represent susceptibility factors for a broad spectrum of neuropsychiatric disorders such as epilepsy, schizophrenia or autism spectrum disorder (ASD) with reduced penetrance and no or rather mild intellectual impairment.[266,272-283,286,299-307] In contrast, biallelic defects in either gene were reported to result in fully penetrant, severe neurodevelopmental disorders.

Strauss et al. reported on a homozygous stop mutation in *CNTNAP2* in Old Order Amish children causing CDFE (Cortical Dysplasia – Focal Epilepsy) syndrome (MIM #610042), characterized by cortical dysplasia and early onset, intractable focal epilepsy leading to language regression, and behavioral and mental deterioration.[184,285] In a former study we reported on homozygous or compound heterozygous defects in *CNTNAP2* or *NRXN1* in four patients with intellectual disability and epilepsy,[308] resembling Pitt-Hopkins syndrome (PTHS, MIM #610954). A possible shared synaptic mechanism that was observed in Drosophila might contribute to the similar clinical phenotypes resulting from both heterozygous and recessive defects in human *CNTNAP2* or *NRXN1*.[308]

To further delineate the clinical phenotype associated with potentially recessive defects in any of the two genes, we screened a group of patients with either severe intellectual disability resembling Pitt-Hopkins syndrome or the phenotypes caused by recessive *CNTNAP2* or *NRXN1* defects. Additionally, we performed mutational testing in patients found to harbor heterozygous deletions in either gene.

MATERIALS AND METHODS

Patients

Our total cohort of patients comprised four different subsets: 1. our new Pitt-Hopkins syndrome-like (PTHSL) screening group, 2. parts of our old PTHSL screening group[308], 3. a group of patients with suspected recessive inheritance, and 4. patients with known heterozygous deletions in one of the two genes. 1. The new PTHSL screening group consisted of 90 patients who were initially referred with suspected Pitt-Hopkins syndrome for diagnostic testing of the underlying gene, *TCF4*, which encodes transcription factor 4. They all had severe intellectual disability and variable additional features reminiscent of the PTHS spectrum such as dysmorphic facial gestalt or breathing anomalies. Mutational testing of *TCF4* revealed normal results. In all of these 90 patients mutational screening of *NRXN1* and *CNTNAP2* was performed in the current study. Molecular karyotyping was performed in 22 of them. This cohort does not overlap with the second subset, our old PTHSL screening group, which is a similar group of 179 patients, reported in a former study.[308] No published information on mutational screening of that group was included in the current study, but previously unpublished information on molecular karyotyping of 23 patients. 3. Nine patients with severe intellectual disability were referred to us specifically for *CNTNAP2/NRXN1* testing because of suspected autosomal-recessive inheritance and/or phenotypic overlap with the previously published patients.[308] 4. In eight patients copy number changes in either *NRXN1* or *CNTNAP2* were identified in other genetic clinics. These were referred to us for mutational screening of the second allele. These patients had variable degrees of intellectual

disability and various other anomalies. An overview on tested patients is given in Table 1. This study was approved by the ethics committee of the Medical Faculty, University of Erlangen-Nuremberg, and written consent was obtained from parents or guardians of the patients.

Table 1 Overview on screened patients

Patient samples used in this Study	Sequencing of *NRXN1* number of patients	Sequencing of *CNTNAP2* number of patients	Molecular karyotyping number of patients
1. new screening sample, n=90	90	90, including C1-C4	22, including N1
2. old screening sample,[308] n=179	published[308], results not used in this study	published[308], results not used in this study	23, not published before
3. specific testing sample*	9	9	
4. *NRXN1/CNTNAP2* deletion group**	5, N2-N6	3, C5-C7	8, (details on arrays see Table 3)

*, Patients were referred to us specifically for *NRXN1/CNTNAP2* testing due to suspected autosomal recessive inheritance and/or phenotypic overlap with the previously published cases. **, Patients were referred to us because of copy number changes in either NRXN1 or CNTNAP2 for screening of the respective second allele.

Molecular karyotyping

Molecular karyotyping was performed in 45 patients without *TCF4* mutation with an Affymetrix 6.0 SNP Array (Affymetrix, Santa Clara, CA), in accordance with the supplier's instructions. Copy-number data were analyzed with the Affymetrix Genotyping Console 3.0.2 software. In patient C3 molecular karyotyping was performed with an Affymetrix 500K array and data analysis was performed using the Affymetrix Genotyping Console 3.0.2 software.
The patients with heterozygous copy number variants (CNVs) referred for sequencing of the second allele, had been tested on different platforms. An overview on the array platforms, validation methods and segregation in the families is given in Tables 2 and 3.

Mutational screening and MLPA

DNA samples of 107 patients were derived from peripheral blood, and if sample material was limited, whole genome amplification was performed using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare, Little Chalfont, Buckinghamshire, United Kingdom) according to the manufacturer's instructions. All coding exons with exon-intron boundaries of CNTNAP2 (NM_014141) and of isoforms alpha1, alpha2 and beta of NRXN1 (NM_004801; NM_001135659; NM_138735) were screened for mutations by unidirectional direct sequencing (ABI BigDye Terminator Sequencing Kit v.3; AppliedBiosystems, Foster City, CA) with the use of an automated capillary sequencer (ABI 3730; Applied Biosystems). Mutations were confirmed with an independent PCR and bidirectional sequencing from

Table 2 Molecular findings in *NRXN1*

NRXN1	Defect	*NRXN1/CNTNAP2* deletion	Validation of Array data	Inheritance	Carrier parent	Other non-polymorphic CNVs	*NRXN1* seq	*CNTNAP2* seq
N1	*NRXN1* deletion of exons 1-4	Affymetrix 6.0 SNP Array chr2:50.860.393-51.208.000 348kb (230 array marker)	MLPA as reported previously[308]	paternal	healthy, normal intelligence	none	no 2nd mutation	normal
N2	*NRXN1* deletion of exons 1-18	Agilent 244K+customized array chr2:50.270.203-51.257.206, 987 kb	customized Oligo array	maternal	learning and behavioral problems	none	no 2nd mutation	normal
N3	*NRXN1* deletion of exons 1-2	Agilent 244A chr2:51.011.745-51.144.527 133kb	qPCR as reported previously[309]	maternal	healthy	21q22.3:44.534.530-44.820.473 pat dup Xp22.33 :0.000.001 -2.710.316 mat dup	no 2nd mutation	normal
N4	*NRXN1* deletion of exons 1-4	Agilent 244A chr2:50.800.974-51.286.171 425kb	FISH with BACs RP11-67N9 and RP11-643L22	paternal	healthy	15q26.1:88.028.337 -88.072.545 mat del 16q12.1:50.773.658 -51.135.179 mat dup	no 2nd mutation	normal
N5	*NRXN1* deletion of exons 3-4	Agilent 244A chr2:50.861.527-51.090.563, 229kb	qPCR as reported previously[309]	paternal	muscular problems, stroke; consanguinity	none	no 2nd mutation	normal
N6	*NRXN1* deletion of exons 1-2	Agilent 244A chr2:51.033.865-51.496.143 462kb	Agilent 244A in parents	*de novo*		none	no 2nd mutation	normal
biallelic defect P3, n=1[308]	*NRXN1* deletion of exons 1-4 + p.S979X	Affymetrix 6.0 SNP Array 113kb		parents heterozygous carriers	healthy			
heteroz. defects with ASD n=18[278,286,299,300,306]	15x *NRXN1* deletion,[286,299,300,306] 2x NRXN1 gain,[300] 1x balanced rearrangement disrupting *NRXN1*[278]	12x Agilent 244K,[299] 3x NimbleGen custom arrays,[300] 1x Affymetrix 100K Array,[286] 1x Affymetrix 10K Array,[306] 66kb-5Mb		6x de novo,[286,299,306] 5x mat,[299,300] 4x pat,[278,299] 3x not available[299,300]		1x duplication 14q24[300]		

mat, maternal; pat, paternal; dup, duplication; del, deletion; ass., associated; FISH, fluorescence in-situ hybridization; BACs, bac clones; oligo, oligonucleotide; qPCR, quantitative Real-Time-PCR; non-polymorphic CNVs: CNVs that have not been reported in the Toronto Database of Genome Variants or have not been identified in one of our molecularly karyotyped healthy control individuals. Seq, sequencing;

Table 3 Molecular findings in CNTNAP2

CNTNAP2	Defect	*NRXN1/CNTNAP2* deletion	Validation of Array data	
C1	*CNTNAP2* c.1175_1176dup; p.D393RfsX51	Affymetrix 6.0 SNP Array, normal results for *CNTNAP2* and *NRXN1*		
C2	*CNTNAP2* c.2153G>A, p.W718X	Affymetrix 6.0 SNP Array, normal results for *CNTNAP2* and *NRXN1*		
C3	*CNTNAP2* c.1083G>A, splice site (p.V361V)	Affymetrix 500K SNP Array, normal results for *CNTNAP2* and *NRXN1*		
C4	*CNTNAP2* c.1083G>A, splice site (p.V361V)	Illumina 317 K SNP Array, normal results for CNTNAP2 and *NRXN1*		
C5	*CNTNAP2* deletion of exons 2-3	Affymetrix 6.0 SNP Array chr7:146.079.333-146.194.785 115kb (69 array marker)	Affymetrix 6.0 SNP Array of the parents	
C6	*CNTNAP2* deletion of exons 3-4	Illumina Human 660W-Quad chr7:146.144.267-146.374.539 230kb (53 array marker)	qPCR as reported previously [310]	
C7	*CNTNAP2* deletion of exons 21-24	Agilent 2X400K chr7:147.702.165-148.378.711 677kb	customized Oligonucleotide array	
published biallelic defects n=13[184,285]	2x *CNTNAP2* deletion of exons 2-9, homozygous,[308] 1x *CNTNAP2* deletion of exons 5-8 + IVS10-1G>T,[308] 10x *CNTNAP2* c.3709delG, homozygous[184,285]	2x Affymetrix 500K / 250K Nsp SNP Array; 1x Affymetrix 6.0 SNP Array,[308] 10x no		
published heterozygous defects n=12[266,272,276,283,284,305]	2x translocation disrupting *CNTNAP2,*[283,284] 1x inversion disrupting *CNTNAP2,*[266] 5x *CNTNAP2* deletion,[272,276,305] 4x missense variant in *CNTNAP2*[266]	3x BAC array,[276] 1x NimbleGen custom array,[305] 220kb-11Mb		

mat, maternal; pat, paternal; dup, duplication; del, deletion; ass., associated; qPCR, quantitative Real-Time-PCR; non-polymorphic CNVs: CNVs that have not been reported in the Toronto Database of Genome Variants or have not been identified in one of our molecularly karyotyped healthy control indivuals

original DNA. Primer pairs and conditions were used as previously described [308]. For splice site prediction, eight different online tools were used as indicated in Table 4. Multiplex Ligation Dependent Probe Amplification (MLPA) for all coding exons of *CNTNAP2* was performed for patients C1-C4 as described previously.[308]

RESULTS

Molecular Testing

Mutational screening of *NRXN1* in 90 *TCF4* mutation negative patients and nine families with suspected recessive inheritance of severe intellectual disability did not reveal any point mutation, while in *CNTNAP2* the heterozygous mutation c.1083G>A in the splice donor site of exon 7 was found in two patients (C3, C4). Eight prediction programs (table 4) showed

	Inheritance	Carrier parent	Other non-polymorphic CNVs	*NRXN1* seq	*CNTNAP2* seq
	paternal	healthy	chr9:9.337.920-10.207.671 mat dup chr13:19.104.340-19.477.398 mat dup	normal	no 2nd mutation; MLPA normal
	not known	not known	none	normal	no 2nd mutation; MLPA normal
	paternal	healthy	none	normal	no 2nd mutation; MLPA normal
	maternal	healthy	pathogenic frameshift mutation in MEF2C (P7)[75]	normal	no 2nd mutation; MLPA normal
	maternal	healthy	none	normal, one silent variant	no 2nd mutation
	maternal	healthy	none	normal	no 2nd mutation
	de novo	healthy	chr7:92.394.428-92.530.356 del chr7:93.464.449-94.430.690 del, both de novo, conventional karyotyping: 46,XX,der(4)t(4;10)(q25;q24), der(7)t(4;7)(q25;q32), der(10)inv(10)(p13q24)(7;10)(q32;p13), de novo	normal	no 2nd mutation
	parents heterozygous carriers				
	2x not reported [276], 4x inherited [266], 2x paternal [272,305], 2x de novo [266,276] 2x balanced in parent (translocation) [283,284]				

diminished splice site recognition for this mutation, which is therefore predicted to result in an in-frame loss of exon 7. This possible splice site mutation was found in one of 384 control chromosomes. Furthermore, in patient C1 the heterozygous frameshift mutation p.D393RfsX51 in exon 8 and in patient C2 the heterozygous stop mutation p.W718X in exon 14 were identified. Due to their nature and location both truncating mutations are predicted to result in mRNA decay and loss of the affected allele. For patient C2 parents were not available, but all other mutations were shown to be inherited from a healthy parent. No defect on the second allele was identified in any of these patients by sequencing and subsequent MLPA-analysis of all coding exons. In 942 controls sequenced by Bakkaloglu et al.,[266] no truncating mutation in *CNTNAP2* was found. No *CNTNAP2* deletion was found in 667 control individuals molecularly karyotyped.[308]

Molecular karyotyping with an Affymetrix 6.0 SNP Array in 45 *TCF4* mutation negative patients revealed a heterozygous deletion within the *NRXN1* gene in one patient (N1). The father was shown to be healthy carrier, and no mutation on the second allele was found in this patient by sequencing of all coding exons.

In three patients with *CNTNAP2* deletions (C5-C7) and in five patients with *NRXN1* deletions (N2-N6) we could not identify any pathogenic mutation on the second allele by sequencing all coding exons. In patient N6 and in patient C7 the deletion within *NRXN1* or *CNTNAP2* was shown to be *de novo*. In all other families the deletion in *CNTNAP2* or *NRXN1* was also identified in one of the parents.

In all patients with a heterozygous defect in *CNTNAP2* we also screened *NRXN1* and vice versa, without observing any anomalies. An overview of localization of novel and published mutations and deletions is shown in Figure 1 and 2. Mutation and array data of novel patients are shown in Tables 2 and 3.

Table 4 Splice site prediction for splice donor variant c.1083G>A

Program	wild type score	mutant score
NNSplice 0.9[311]	0.99	0.6
HSF V2.4[312]	91.56	80.98
MaxEntScan[313]		
Maximum Entropy Model	8.37	3.38
Maximum Dependence Decomposition Model	11.88	9.78
First-order Markov Model	7.5	3.88
Weight Matrix Model	8.9	5.73
Splice Site Score Calculation[275]	8.1	5.2
Splice Site Analyzer-Tool[314]	83.27 ∆G -7.1	71.36 ∆G -4
Splice Predictor[315]	0.967	splice site not recognized
NetGene2[316]	0.95	0.55
SplicePort[276]	1.06619	0.26169

Clinical Findings

Four of six patients with heterozygous CNVs in *NRXN1* were severely intellectually disabled (N1-N4). Three had epilepsy and one episodic hyperbreathing. Patients N5 and N6 were only mildly intellectually disabled and N5 additionally had various malformations like choanal atresia, anal atresia, and skeletal anomalies. All patients had absent or impaired language abilities, while motor development was normal or only mildly delayed in four of them. The deletion in patient N6 was shown to be de novo, in all other families one parent was shown to be carrier of the deletion. The mother of N2 was reported to have had learning difficulties, all others were reported to be healthy and of normal intelligence. However, detailed neuropsychiatric testing was not performed. Summarized clinical details of the patients are shown in Table 5.

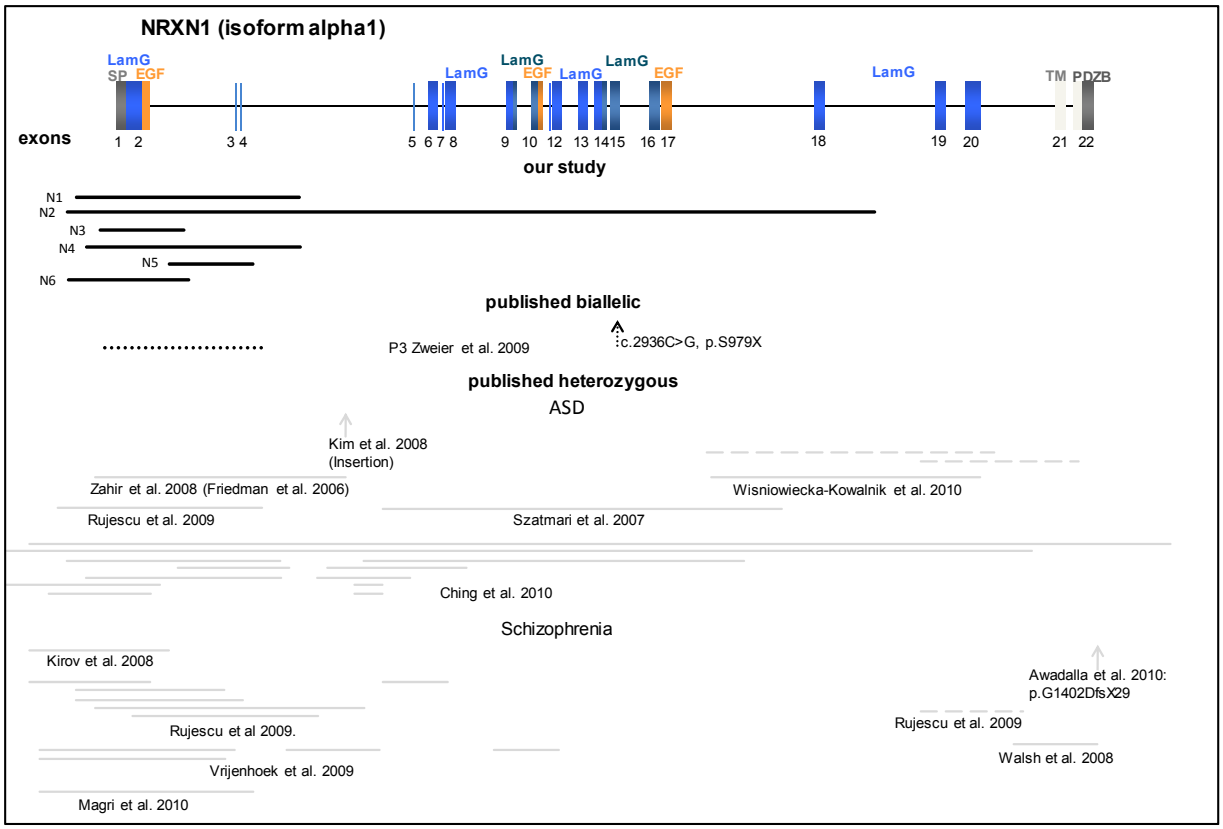

Figure 1 Schematic drawing of *NRXN1* with localization of novel and published mutations and deletions.

Schematic drawing of genomic structure of alpha 1 isoform of *NRXN1* showing domain-coding exons and localization of mutations and deletions. Deletions found in our study are represented by black bars. Published biallelic aberrations are shown with black dotted lines, whereas heterozygous losses and gains are marked by grey solid and dashed lines, respectively.

Abbreviations are as follows: SP, signal peptide; LamG, laminin-G domain; EGF, epidermal growth factor like domain; TM, transmembrane region; PDZBD, PDZ-domain binding site.

All seven patients with heterozygous defects in *CNTNAP2* in this study showed severe to profound intellectual disability. Speech was lacking in four patients (C1, C4-C6) and reported to be simple in C7. Patient C3 lost her speech ability at age 2.5 years. Motor impairment was also severe with no walking abilities in three patients (C4-C6), patient C7 started to walk at the age of 15 months, and patients C1 and C3 lost this function at age 2.5 – 3 years. Five patients had seizures. As far as data were available, epilepsy was of early onset and difficult to treat. At least in two of the patients episodes of hyperbreathing were reported. Congenital anomalies and malformations such as tetralogy of Fallot, pyloric stenosis, and variable other anomalies or septo-optical dysplasia were reported in patients C1 and C5, respectively. In the parents shown to be carriers, no neuropsychiatric anomalies were reported. However,

detailed neuropsychiatric testing was not performed. Summarized clinical details of the patients are shown in Table 6.

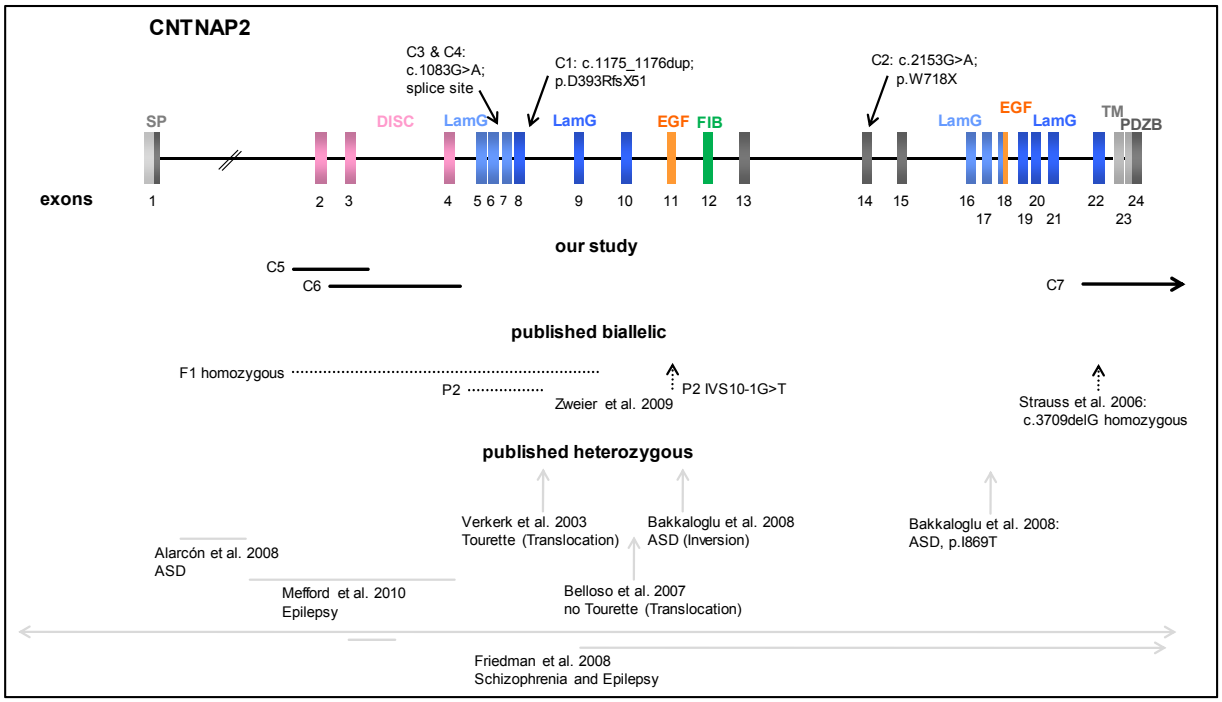

Figure 2 Schematic drawing of *CNTNAP2* with localization of novel and published mutations and deletions.

Schematic drawing of genomic structure of *CNTNAP2* showing domain-coding exons and localization of mutations and deletions. Mutations and deletions found in our study are represented by black arrows and bars. Published biallelic aberrations are shown with black dotted lines, whereas heterozygous defects are shown in grey. Abbreviations are as follows: SP, signal peptide; DISC, discoidin-like domain; LamG, laminin-G domain; EGF, epidermal growth factor like domain; FIB, fibrinogen-like domain; TM, transmembrane region; PDZBD, PDZ-domain binding site.

DISCUSSION

NRXN1. While the majority of the novel patients had severe intellectual disability, only two of the patients, N5 and N6, with heterozygous deletions in NRXN1 had mild intellectual disability as reported before for this kind of defects.[278,280,286,299,300] Additionally, patient N5 had various congenital malformations and anomalies. Interestingly, one recently published patient with a NRXN1 defect and no significant intellectual impairment was reported with similar malformations resembling the VACTERL spectrum.[299] Mild skeletal anomalies were also reported in the patient published by Zahir et al..[286] A larger number of patients and therefore further delineation of the phenotype will probably clarify a possible relation of such malformations to NRXN1 defects.

Table 5 Clinical findings associated with defects in *NRXN1*

NRXN1	Sex & Age	ID	Speech	Age of Walking	Seizures age of onset	birth Weight, Heigth, OFC	Weight Height OFC	Behavioral anomalies/ Stereotypies	Facial dysmorphism	Other findings
N1	m, 14y	Severe	3y max. 10 words, lost this function	14mo	yes	2900 g 52 cm 34 cm	P25-P50 P25-P50 P90	yes, puts objects in his mouth	large mouth, widely spaced teeth, UPF, strabism	hyperbreathing
N2	m, 6y	Severe	24mo single words, two word comb. receptive > expressive	16mo	none	3740 g 51 cm 38.5 cm	Normal <P3 >P95	none	macrocephaly (also maternal and paternal), large mouth, retrogenia	muscular hypotonia, MRI: wide ventricles
N3	m, 3y 4mo	Severe	no active speech	14mo	none	3350 g 52 cm 35 cm	P50-P75 P75-P90 P50-P75	yes	none	none
N4	f, 16y	Severe	None	No	grand mal 4y	3530 g 51 cm 33 cm	P10-P25 P25-P50 <P5	yes, hand licking	broad nasal tip, pointed chin	drooling, friendly
N5	m, 21y	Mild	Impaired	not known	grand mal, 6y (until age 11y)	3300 g 51 cm 33 cm	P3-P10 <P3 P50	none	mild facial asymmetry, small ears, broad nose, broad mouth, bushy eye brows, high arched palate, cleft lip	pectus excavatum, single transverse palmar crease, choanal atresia, anal atresia, thick finger joints, ureterstenosis, delayed bone age, spondyloptosis L5/S1
N6	f, 6y 3mo	Mild	2 y: first words, mainly active speech affected	21mo	none	2820 g 50 cm 35 cm	P10-P25 P3 P10-P25	none	protruding ears	muscular hypotonia (improved), scapulae alatae, mild lordosis, tendency to diarrhea
biallelic defect P3, N=1[308]	f, 18y	Severe	None	2y	none	3450 g normal	P50-P75 P50-P75 P25	yes, hypermotoric behavior	broad mouth, strabism, protruding tongue	excessive drooling, developmental regression, abnormal sleep-wake-cycles, decreased deep-tendon reflexes upper extremities, hyperbreathing
published heteroz. defects ass. with ASD N=18[278,286,299,300,306]		7x normal,[299] 3x learn. problems,[299,300] 2x dev. Delay,[299,306] 3x mild ID,[278,286,300] 2x moderate ID[299]	14x language delay[286,299,300,306]	5x motor delay[286,299]	1x yes[299]	not reported	not reported	11x ASD or Asperger syndrome[278,286,299,300,306]	11x mild dysmorphic features[286,299,300]	1x VACTERL association,[299] 1x mild skeletal anomalies,[286] 4x hypotonia, 2x ventricular septum defect, 3x hemangioma[299]

TOF, tetralogy of Fallot; f, female; m, male; y, year; mo, months; ASD, autism spectrum disorder; published reports on CNTNAP2 and NRXN1: only papers containing clinical data are cited; ass., associated; P, centile; learn., learning problems.

Table 6 Clinical findings associated with defects in *CNTNAP2*

CNTNAP2	Sex & Age	ID	Speech	Age of Walking	Seizures age of onset	birth Weight Heigt OFC	Weight Height OFC	Behavioral anomalies/ Stereotypies	Facial dysmorphisms	Other findings
C1	f, 8y	Severe	none	2y with aid, lost this function (3y)	yes, resist. to treatment	2430 g 45 cm not reported	<P3 <P3 <P3	hand movements	synophrys, long eyelashes, prominent columella, short philtrum, arched palate, widely spaced teeth, prominent jaw	happy, affectionate, TOF, pyloric stenosis, vesicoureteric reflux, agenesis of labia minora, hirsutism, tapering fingers
C2	m, 18y	Severe	?	?	complex, early onset	?	?	?		hyperbreathing, apnoe episodes
C3	f, 11y	Severe	few words, lost this function	2,5y, lost this function	3y	3510 g	P10 <P3 P10	yes	broad mouth, protruding tongue	developmental regression from 15mo, swallowing problems, nocturnal laughing, scoliosis, spastic tetraparesis, hyperreflexia, constipation,
C4[75]	f, 7y	Profound	none	none	3-6mo	3400 g	P5 <P2 P50	yes	broad forehead, prominent ear lobes, widely spaced teeth, tented upper lip	exotropia, heterochromasia, high pain threshold, cold feet, sleeping problems, joint hyperlaxity
C5	f, 2y 8mo	Profound	none	none	none	4030 g 53 cm 38 cm	P75 P25-50		high arched palate, upslanting palpebral fissures, small teeth, prominent forehead	septo-optical dysplasia, MRI: agenesis of septum pellucidum

C6	8y	Profound	none	no	yes, resist. to treatment	1160 g 35 cm 28 cm	<P3 <P3 <P5		mild synophrys, low set, large ears, fleshy ear lobes, thin upper lip, low frontal hairline	birth at 29th week of gestation, blindness, hydrocephalus, ductus arteriosus, syndactyly toes 2-3, hypotonia, spasticity of legs, obstipation, liquid uptake by PEG tube
C7	f, 8y	moderate to severe	simple	15mo	none	3860 g 54 cm 34 cm	P25-P50 P50 <P5	suspected in infancy	epicanthal folds, tented upper lip, short columella, bulbous nose	overfriendliness, pubertas praecox, delayed bone age, retentive memory, excessive empathy, autoagressive behavior, flat feet
published biallelic defects N=13[184,285]	2x f, 1x m, 10x not reported,1-20y	Severe	2x no, 1x single words,[308] 10x yes, but regression[184,285]	2x normal, 1x not known,[308] 10x 16mo-30mo[184,285]	13x yes, 4mo-30mo	not reported	<P3-normal not reported <P3-P99	8x yes,[184,308] 1x tooth grinding and repetitive hand movements[308]	2x wide mouth and thick lips[308]	1x dry skin, 1x regression, 1x cerebellar hypoplasia, 3x hyperbreathing,[308] 10x developmental regression with onset of seizures, 9x decreased deep tendon reflexes,[184,285] 4x MRI: cortical dysplasia,[184] 1x MRI: leukomalacia, 1x hepatosplenomegaly[285]
published heterozygous defects N=12[266,272,276,283,284,305]		6x not reported,[266,272,305] 1x normal,[276] 2x mild-moderate,[266,276] 3x severe[276,283,284]	6x not reported,[266,272,305] 1x normal,[276] 3x speech impairment,[276,283] 2x no[276,284]	11x not reported,[266,272,276,283,305] 1x no[284]	5x not reported,[266,272] 2x no,[283,284] 5x yes,[266,276,305] 0y-34y	not reported	not reported	8x yes[266,272,276]	not reported	1x multiple congenital malformations,[284] 1x Gilles de la Tourette syndrome,[283] 3x Schizophrenia[276]

TOF, tetralogy of fallot; f, female; m, male; y, year; mo, month; ASD, autism spectrum disorder; published reports on CNTNAP2 and NRXN1: only papers containing clinical data are cited; ass., associated; P, centile

All other four patients with heterozygous NRXN1 deletions were severely intellectually disabled without specific further anomalies. Their phenotype resembled the patient reported with a compound heterozygous defect in this gene.[308] Except for patient N4, speech impairment was severe compared to a rather mild motor delay. Because of the severe phenotype in the patients in contrast to the normal or only mildly impaired intellectual function in the respective carrier parent, a defect of the second allele was suspected in the patients, but not found.

CNTNAP2. Most of the clinical aspects and the severity of intellectual disability in the herewith reported patients with heterozygous CNTNAP2 defects resembled those observed in patients with biallelic defects in CNTNAP2 reported before (Table 6). Two of the patients (C1, C3) showed language and motor regression correlating with onset of epilepsy. All others showed lacking or severely impaired speech development. However, in contrast to the published patients with recessive defects and normal or only mildly delayed motor development,[184,308] all but one patients in this study also showed severe motor retardation. We could not identify a defect on the second allele in any of the novel patients. In most of the families the defect was inherited from a healthy parent. Despite a significantly higher frequency ($p<0.01$, Fisher's exact test) of two truncating mutations in our cohort of 99 severely to profoundly intellectually disabled patients compared to no truncating mutation in 942 normal controls[266] definite proof that the respective mutation is fully responsible for the phenotype is so far lacking. This also applies to the other identified defects in CNTNAP2 or NRXN1.

Congenital malformations as described in patients C1 or C5 (Table 6) have not yet been reported in any other patient with a CNTNAP2 defect. Furthermore, the fact that the expression of the gene is restricted to the nervous system[297] does not explain these anomalies. Therefore, another genetic cause for these malformations might exist. Thus it is difficult to define if the intellectual disability is associated with the CNTNAP2 mutation at all in these patients. Other factors like premature complicated birth in patient C6 might contribute to impaired intellectual function. C3 and C4 carried the same splice site mutation and both showed a similar phenotype with severe intellectual disability and seizures, C3 also with breathing anomalies. In a parallel research project, a mutation in the MEF2C gene was identified in patient C4 and shown to be capable of causing all of her symptoms.[75] Therefore, it remains unclear if this splice mutation has a pathogenic effect at all, or only a mild effect that is masked by the severe consequences of the MEF2C mutation. The fact that this variant is supposed to lead to an in-frame loss of a single exon with a possibly milder effect than more deleterious defects supports the idea of no or only minor relevance of this splice mutation. Regarding the relatively high frequency of the splice site mutation in two families

and one control, a founder effect might be considered, however, common regional background in these persons is not obvious.

Expanding the observations from previous studies we now found that heterozygous defects in CNTNAP2 or NRXN1 can also be seen in association with severe intellectual disability. Possible explanations might be: 1. No pathogenic relevance of the identified defect. This might indeed be the case for those patients with a "mild mutation" such as the splice-site mutation in CNTNAP2, or for patients with an atypical phenotype or congenital malformations. In those, the true causative defect might not be detected yet. However, published data and our data together still support a pathogenic role for both genes in neurodevelopmental disorders. 2. Inability to identify a defect on the second allele in spite of extensive screening for mutations and/or deletions. However, mutations in regulatory elements or in additional alternative isoforms cannot be excluded in any case. 3. A larger phenotypic variability associated with heterozygous defects in each gene. The finding of homozygous or compound heterozygous defects in previous patients with severe phenotypes[184,285,308] indicates the existence of second hits or additional major contributors. These might not necessarily be affecting the same gene. Only recently, a two-hit model for severe developmental delay in patients with a recurrent 16p12.1 microdeletion was postulated.[45] This might also be the case for microdeletions or even point mutations within a single gene as already reported for digenic inheritance in specific ciliopathies like Bardet-Biedl syndrome.[48] In four of our patients additional de novo or parentally inherited CNVs were identified (see Tables 2 and 3), however, the significance of these CNVs is unclear. The possible functional synaptic link between CNTNAP2 and NRXN1[184,285,308] prompted us to screen CNTNAP2 in patients with NRXN1 defects and vice versa, however, without any mutation detected.

CONCLUSION

We found heterozygous defects in CNTNAP2 and NRXN1 in patients with severe intellectual disability, therefore expanding the clinical spectrum associated with monoallelic defects in either gene. This large variability implicates difficulties for genetic counseling in such families. To explain the larger phenotypic variability and severity in some patients we suggest a contribution of major additional genetic factors. To identify these possible contributors and modifiers will be a great challenge for the near future.

AUTHORS' CONTRIBUTIONS

BA, IB, EKB, DH, JH, JKl, IM, EP, ST, EW, and GW acquired and provided clinical data and samples of their patients. AG, ABE, HE, KH, JKo, SN, RU, ARe, and CZ created and

analyzed the molecular data. ARe and ARa revised the manuscript critically for important intellectual content. CZ designed and supervised the project, and together with AG drafted the manuscript. All authors read and approved the manuscript.

ACKNOWLEDGEMENTS

We thank the contributing clinicians, the patients and their families for participating. We thank Christine Zeck-Papp for excellent technical assistance and Dr. D. Müller and Dr. A. Kobelt for providing clinical details. This study was funded by a grant from the DFG (ZW184/1-1) and by the German MR-NET funded by the BMBF.

Chapter 6

Genetics of ID – an inventory

Christiane Zweier, Korinna Kochinke, Shivakumar Keerthikumar, Bonnie Nijhof, Misa Fenckova, Martin Oti, Tjitske Kleefstra, Martijn Huynen, Annette Schenck

part of a larger project – in preparation

SUMMARY

Intellectual disability (ID) is clinically and genetically highly heterogeneous, and the underlying genes are only incompletely identified so far. Understanding individual ID gene function and connecting genes and proteins in common functional networks and complexes will be a prerequisite to better understand the mechanisms of cognitive function and dysfunction and to possibly develop future therapeutic interventions.

To facilitate both disease gene prediction as well as functional prediction, we now established a curated list of all known, highly reliable ID genes and introduced them into a database that is supplemented with many other datasets. We furthermore established a classification of ID genes based on the clinical manifestation of the associated disorders and analyzed whether and how the genes in the ID classes related to clinical and biologic features such as inheritance, expression, synaptic localization or haploinsufficiency.

INTRODUCTION

Intellectual disability (ID) affects 2%-3% of the population in western countries and is characterized by significant limitations in both intellectual functioning and adaptive behavior that begins before the age of 18 years and is reflected in an IQ below 70.[1,2] ID is clinically and genetically extremely heterogeneous. It can occur isolated in non-syndromic ID or in the context of syndromes with additional specific features such as malformations or other anomalies, specific biochemical findings, or a distinct facial phenotype. Though an increasing number of underlying genetic alterations for ID has been identified, the etiology still remains unsolved in many patients.[4,14]

Next to chromosomal aberrations and to some complex causes, that are yet poorly understood, a large fraction of ID disorders is supposed to be based on monogenic defects.[317] These are assumed to be highly heterogeneous. In contrast to chromosomal aberrations, that can now reliably be detected with molecular karyotyping,[10,15,70] the systematic and large-scale detection and identification of disease associated genes and of point mutations remains still challenging and probably contributes to the still high proportion of patients with unclear cause of ID. Due to the advances in disease gene identification over the past decade, more than 400 genes involved in ID have been identified.[3] Estimations, based on the number of known X-chromosomal ID genes, assume a total of 1500 to 2000 ID-genes.[2,3]

Since recently, next-generation-sequencing (NGS) technologies provide for the first time a possibility to systematically screen for the underlying mutations, both in X-linked recessive (Kalscheuer et al., Next-generation sequencing in 248 families with X-linked intellectual disability; Abstract #84, Presented at the 12th International Congress of Human

Genetics/61st Annual Meeting of The American Society of Human Genetics, October 13th, 2011, Montreal, Canada) and autosomal recessive ID[63] as well as in sporadic ID by identifying *de novo* mutations.[34,35,42] NGS can be utilized for the diagnostic detection of mutations in known ID genes as well as for the identification of novel ID genes.

However, the technical advance is accompanied by several challenges. Mutations in novel candidate genes have to be confirmed, either by finding more patients, or by proving a pathogenic relevance of the mutation or the candidate gene functionally. In order to be able to predict ID genes, we first have to understand the specifics of an ID gene. This we can best deduce from the features of the currently known ID genes. Furthermore, molecular networks have to be established to provide the basis for the long-term creation of therapeutic intervention which will rather target central points in pathways than single genes.

Linking ID genes in common molecular or functional pathways indicates a general regulatory role of these networks in cognitive function and dysfunction[2,318-321] and supports the idea that similar phenotypes are caused be defects in functionally related genes.[178] Examples for such networks are “rasopathies”,[179] “ciliopathies”, “cohesinopathies” or “channelopathies”[180], all comprising a group of identical or overlapping clinical phenotypes caused by defects in genes belonging to the same pathway or biological process.

To cope with the arising challenges and to further identify and delineate the complex networks of cognitive function and dysfunction, a systematic, large-scale and collaborative interdisciplinary approach is required.

Here we present a comprehensive inventory of all known ID associated genes compiled into a database together with available datasets on functional, molecular and clinical properties. These highly reliable disease genes are classified according to their phenotypic outcome when mutated. By investigating correlations among molecular, functional, and clinical datasets, patterns for future disease gene prediction and delineation of networks and pathways are established.

METHODS

Gene list

A list of genes in which mutations are sufficient to cause ID (ID genes) was compiled from primary and secondary[257] literature and public databases like OMIM (http://www.ncbi.nlm.nih.gov/omim/), using text-mining and search terms such as “mental retardation”, “intellectual disability”, “cognitive disability”, or “developmental delay”.

This list has been subsequently manually curated in order to comprise only genes with reliable association to ID. Reasons to exclude genes were: 1. Low evidence, indicated by only a single patient with a mutation in a particular gene, by gene-disrupting translocations

or deletions but no mutational confirmation of the candidate gene in other patients, or by clinical description of the disorder without genetic testing or confirmation. 2. Pure neurodegenerative manifestation, indicated by secondary onset of intellectual disability with regression of initial normal abilities, or by a progressive disease course with deterioration of cognitive abilities. 3. Very early lethality, thus precluding proper assessment of psychomotor development. 4. Treatability, indicated by avoidance of decline in cognitive ability by substitution of certain factors, e.g. consequences of hypothyroidism which can be avoided by thyroid hormone substitution. 5. Neurologic phenotype without clear indication of cognitive impairment.

Phenotypic Classification

All high confident ID genes and their ID associated disease phenotypes were classified into nine main classes, according to the occurrence of non-syndromic or syndromic ID with or without congenital malformations (syndromic and non-syndromic groups, x-axis) and according to character, severity, and penetrance of ID (classic severe, mild, and non-classic groups, y-axis) (figure 1).

For the phenotypic classification clinical information from the respective entries in Gene reviews (http://www.ncbi.nlm.nih.gov/sites/GeneTests/review?db=GeneTests) and/or OMIM (http://www.ncbi.nlm.nih.gov/omim) were used. Where necessary, information from primary literature was consulted. In case of genetically heterogeneous disorders only gene-specific clinical information was used for the respective phenotype classification.

The group of “classic” ID (classes 1-6) comprises ID disorders characterized by primary ID that is stable and without obvious signs of regression or degeneration. They are further distinguished into the classic severe group (CS) (classes 1-3), characterized by moderate to profound ID with full penetrance and in the classic-mild-to-moderate-group (CM) (classes 4-6) with either very variable ID or ID in the mild/borderline range. The “non-classic” ID group (NC) (classes 7-9) comprises ID disorders which are predominantly characterized by non-ID clinical features and in which ID occurs only in a low frequency or disorders in which ID is atypical in its manifestation, for example progressive. Regarding class 8 this resulted in a splitting into 8a for rare ID and 8b for atypical course of ID.

The syndromic classes (1, 2, 4, 5, 7, 8) contain disorders that show recognizable or specific phenotypes such as malformations, facial dysmorphism or specific biochemical/metabolic anomalies. In syndromic classes 1, 4, and 7, structural anomalies of limbs, brain or other organs are reported (SWSM, syndromic with structural malformations), in syndromic classes 2, 5, and 8 ID is accompanied by a specific syndromic feature but not by structural malformations (SWOSM, syndromic without structural malformation). The non-

syndromic group (NS) comprises classes 3, 6, and 9 and includes ID disorders without any additional recognizable phenotype.

The phenotypic classification was done by the same clinical expert and the main classes were revised independently by a second clinician. Discrepancies were discussed and jointly agreed on.

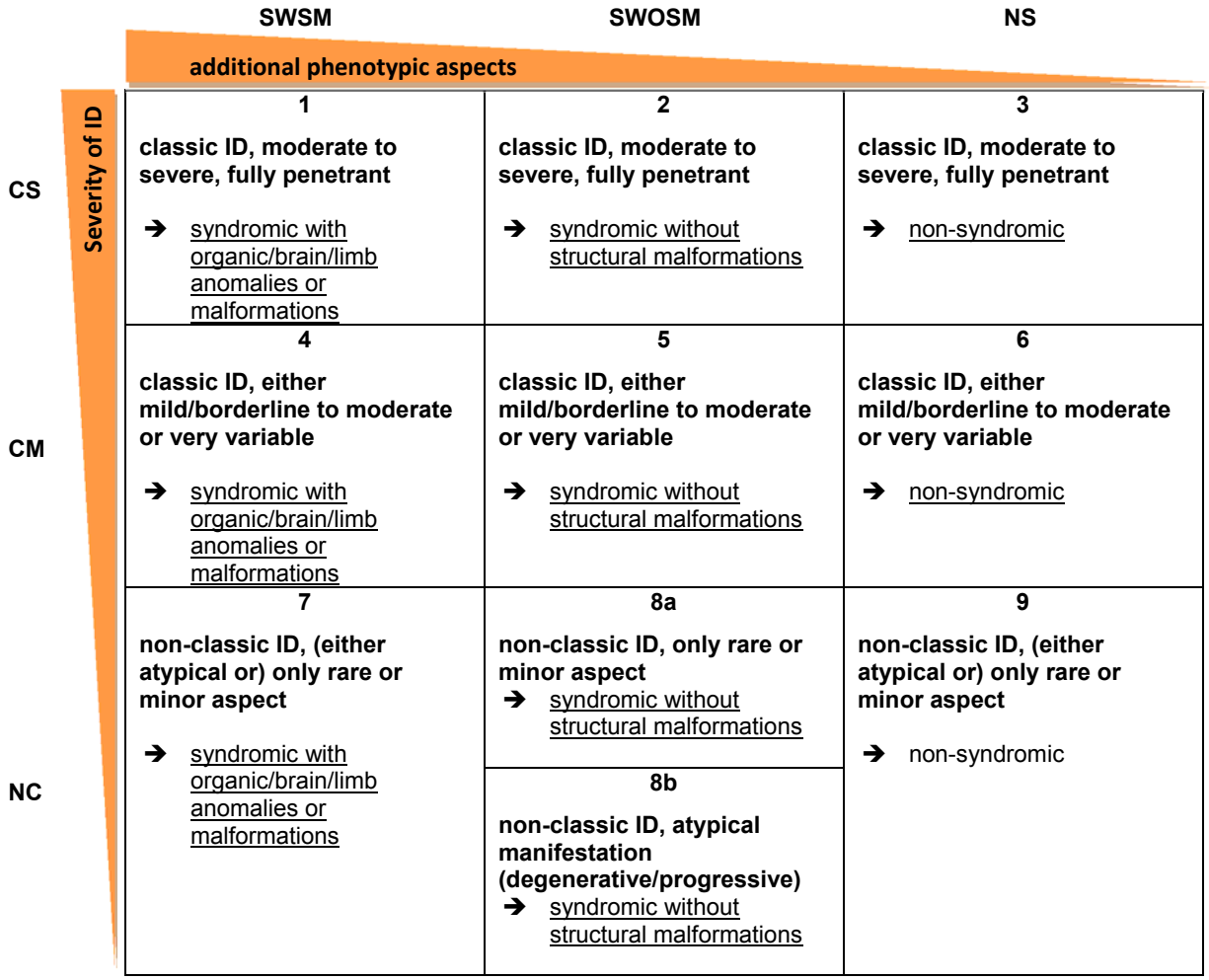

Figure 1 Definition of a novel ID classification scheme: Nine individual clinical classes organized in six groups.

The SWSM group contains disorders that are syndromic with structural malformations and comprises classes 1, 4, and 7; the SWOSM group contains disorders that are syndromic without structural malformations and are included in classes 2, 5, and 8; the NS group contains non-syndromic disorders, that are included in classes 3, 6, and 9. The CS group contains disorders with classic severe manifestation of ID and comprises classes 1, 2, and 3; the CM group includes disorders with classic mild to moderate or variable manifestation of ID and comprises classes 4, 5, and 6; the NC group contains disorders with a rare or atypical manifestation of ID and includes classes 7, 8, and 9.

In addition to the classification system with nine main classes we established an additional classification system for ID-accompanying phenotypes. These comprise 25 additional features describing further symptoms and anomalies of various organ systems (table 1). Letters A-X indicate the presence of specific clinical features and were added when the (estimated) reported frequency of the respective symptom was around 20 to 30%. Letter Z indicates limited clinical data, usually due to a small number of affected patients.

Table 1 Summary of additional phenotypic information

Letter	feature(s) or organ systems
A	short stature/growth failure; AA, short stature syndrome, Aa (acquired), Ac (congenital)
B	Microcephaly; BB, microcephaly syndrome; Ba (acquired), Bc (congenital)
C	early lethality or shortened life span
E	epilepsy/seizures
F	overgrowth/tall stature and/or macrocephaly
G	developmental regression and/or disease progression
H	neurological symptoms, e.g. spasticity, ataxia, severe hypotonia, etc.
I	cancer, tumors
J	immunological anomalies (for example susceptibility to infections)
K	endocrine anomalies
L	brain malformations or (specific) brain anomalies; structural malformations (in groups 1,4,7) as well as non-structural MRI anomalies (e.g. white matter anomalies, myelination anomalies in groups 2,5,8)
M	metabolic, mitochondrial
N	obesity
O	vegetative symptoms (breathing anomalies, obstipation, sweating, sleeping, etc.)
P	behavioral anomalies, autism, autistic behavior, stereotypies, aggression
Q	muscular anomalies, including cardiomyopathies
R	blood anomalies, e.g. anemia or coagulation defects
S	skin, hair, nails
T	eye anomalies (structural malformations, but also minor anomalies like myopia)
U	bone/skeletal anomalies; Ua: limb malformations/anomalies (syndactyly, polydactyly, split hand/foot, etc.); Ub: skeletal/skull malformation/anomalies (fused vertebras, craniosynostosis, etc.); Uc: cleft palate, midline clefts
V	congenital heart defects, cardiac malformations
W	urogenital anomalies and/or renal malformations/disease
X	other malformations
Z	only few single patients (<5) or only 1 family, limited clinical data

SysID database

All manually curated ID genes and their classification were implemented into a MYSQL database and supplemented with a variety of additional information and details on different levels. Gene related details include a short gene description, the official gene number, synonyms, and chromosomal location. Gene related disease information is given by the associated diseases, both ID- and non-ID-related ones, if applicable, by MIM numbers of

those (http://www.ncbi.nlm.nih.gov/omim/), and by mode of inheritance. Further clinical information is provided either by a non-standardized summary of characteristic symptoms or by referring to PMIDs from GeneTest review entries (http://www.ncbi.nlm.nih.gov/sites/GeneTests/review?db=GeneTests) or to a reference from primary literature.

Datasets for enrichment analyses

Datasets used for enrichment analyses were: 1) A dataset of 1458 human postsynaptic density (hPSD) proteins that were identified from human neocortex by mass spectrometry[322] and that constitute 7% of the human genome. 2) A dataset of 299 haploinsufficiency genes (HI) that were identified by text-searching and database-mining strategies[323] and which constitute 1,5% of the human genome. 3) A microarray dataset of human gene expression data (GNF atlas) containing expression profiles from 84 different human tissues and cell types.[324] The significance of the gene expression was statistically tested using a Wilcoxon two sample test (p-value >0.05).

Enrichment analyses

All analyses were done on an initial set of 388 ID genes that were reported and classified in 2010. Enrichment analyses were performed for the above mentioned hPSD and HI datasets as well as for inheritance patterns. Two different backgrounds were used: 1) The human genome, containing ~20500 genes, and 2) the background of all 388 listed ID genes. The human genome background was used to determine the characteristic features of ID genes in general. The 388 ID gene background was used to determine features of specific ID gene classes among all ID genes. To determine enrichment of inheritance patterns, only the 388 gene background was used.

The genes from each of the six groups (CS, CM, NC, SWSM, SWOSM, NS) or each of the nine individual classes (class 1-9) were mapped to the respective dataset, and the mapping frequency was determined (number of mapped genes/number of genes in group or class). After establishing the frequency of genes from the respective dataset to the background data (either: all genes from the dataset/human genome background of ~20500 genes, or: mapped genes from the dataset/background of 388 genes from the list), the fold enrichment was calculated (frequency of mapped genes in groups or classes/frequency of mapped genes in the background). Uncorrected P-values were determined with a chi-square test.

Expression analyses

Expression levels for various tissues from the GNF atlas[324] (Genomics Institute of the Novartis Research Foundation) were mapped to the genes from the gene list. Subsequently, the average expression level for each tissue was determined for each of the six groups and the nine individual classes and compared to the mean expression level of either the human genome background or the background of the 388 gene list. The significance of gene expression was statistically tested using the Wilcoxon two sample test (p-value < 0.05).

RESULTS

Classes

In May 2013 the list contained 518 genes (data not shown) that were distributed over nine main classes according to their phenotypes (table 2). Table 3 shows the distribution for the earlier 388 gene set (supplement) that was used for the enrichment analyses.

Table 2 current set of 518 genes

	SWSM - 150 - 28%	**SWOSM - 333 - 62%**	**NS - 56 - 10%**
CS - 150 - 27%	class 1 - 55 - 10%	class 2 - 89 - 15%	class 3 - 10 - 2%
CM - 266 - 46%	class 4 - 80 - 14%	class 5 - 147 - 25%	class 6 - 46 - 8%
NC - 150 - 27%	class 7 - 26 - 5%	class 8 - 123 - 21% subclass 8a - 50 subclass 8b - 79	class 9 - 1 - 0.2%

Of the 518 genes from current gene list, 459 belonged to only one clinical ID class, while 55 genes belonged to two different classes, and two genes to three or four classes, respectively. Additional to the classified ID phenotype, mutations in 93 of the 518 genes (18%) also cause non-ID disorders with a different OMIM disease number, e.g. a muscular or skeletal disorder without ID.

Table 3 earlier set of 388 genes

	SWSM - 106 - 26%	**SWOSM - 256 - 63%**	**NS - 45 - 11%**
CS - 102 - 24%	class 1 - 36 - 8%	class 2 - 60 - 14%	class 3 - 10 - 2%
CM - 200 - 47%	class 4 - 59 - 13%	class 5 - 112 - 26%	class 6 - 35 - 8%
NC - 126 - 29%	class 7 - 21 - 5%	class 8 - 104 - 24% class 8a - 46 class 8b - 64	class 9 - 1 - 0.2%

Inheritance

The 388 ID gene set (2010) comprises 227 genes (59%) with autosomal recessive mutations, 89 X-linked genes (23%), 80 genes with autosomal dominant mutations (21%), and nine mitochondrially encoded ID genes (2.3%).

The 518 ID gene set from 2013 included 306 genes (59%) with autosomal recessive mutations, 102 X-linked genes (20%), 125 genes with autosomal dominant mutations (24%), and nine mitochondrially encoded ID genes (1.7%).

Enrichment analyses

Enrichment and distribution of post-synaptic density localization within ID genes

ID gene enrichment analysis for a set of 1458 post-synaptic density genes[322] was performed against the background of the human genome, containing ~20500 genes. When comparing the CS, CM, and NC groups, they were significantly, but similarly enriched for hPSD genes. Differential enrichment was found among the syndromic versus non-syndromic groups. Non-syndromic (NS) genes showed 4.1 fold enrichment compared to only 2.4 fold enrichment of the syndromic genes. Interestingly, among the syndromic genes, the SWSM genes did not show enrichment for hPSD genes at all. Accordingly, among the individual classes, class 3 and class 6 showed highest enrichment, 5.6 fold and 3.6 fold, respectively. Classes 1, 2, 5, and 8 were lower, but still significantly enriched for hPSD genes, too (table 4).

Performing the enrichment analysis against the background of 388 ID genes did, due to less statistic power, not result in significant differences of hPSD enrichments between groups or classes of ID genes (data not shown).

Enrichment and distribution of haploinsufficiency within ID genes

ID gene enrichment analysis for a set of 299 haploinsufficiency genes[323] was performed against the background of the human genome, containing ~20500 genes. Both classic (CS, CM) and non-classic (NC) groups showed significant enrichment for HI genes. Regarding syndromic (SWSM, SWOSM) and non-syndromic groups (NS), significant enrichment was found among all groups, too. However, with a very high 12.3 fold enrichment for HI genes in the SWSM group, and a 7.6 fold enrichment among syndromic groups as a whole, haploinsufficiency is less enriched among non-syndromic groups (4.6 fold). Accordingly, class 1 (11.4 fold), class 4 (11.6 fold), and class 7 (13.1 fold) showed highest enrichment for HI ID genes. Also class 5 (8 fold) and class 8a (11.9 fold) showed significant enrichment (table 4).

Performing the enrichment analyses against the background of all ID genes did not result in significant enrichment for HI genes in any of the groups or classes.

Enrichment for inheritance patterns within ID genes

ID gene enrichment analysis was performed with three major inheritances modes (autosomal recessive, autosomal dominant, X-linked (excluding X-linked dominant)) against the background of the 388 ID gene set.

Table 4 Enrichment analyses for hPSD, HI, and inheritance

	classic and non-classic groups		
	CS – 102		
hPSD	**21, 2.9 fold, p = 4.04e-07**		
HI	**7, 4.7 fold, p = 4.30e-05**		
aut dom	21, 1 fold, p = 1		
aut rec	60, 1 fold, p = 1		
x-linked	25, 1.2 fold, p = 0.4		
	CM – 200		
hPSD	**37, 2.6 fold, p = 1.412e-09**		
HI	**25, 8.6 fold, p = 2.08e-34**		
aut dom	42, 1 fold, p = 1		
aut rec	114, 1 fold, p = 0.8		
x-linked	50, 1.2 fold, p = 0.2		
	NC – 126		
hPSD	**20, 2.23 fold, p = 0.0003**		
HI	**14, 7.6 fold, p = 2.44e-17**		
aut dom	25, 1 fold, p = 1		
aut rec	75, 1 fold, p = 0.9		
x-linked	16, 0.6 fold, p = 0.7		
	syndromic and non-syndromic groups		
	SWSM – 106	**SWOSM – 256**	**NS – 45**
hPSD	12, 1.6 fold, p = 0.14	**50, 2.8 fold, p = 7.06e-14**	**13, 4.1 fold, p = 7.9e-08**
HI	**19, 12.3 fold, p = 1.7e-40**	**21, 5.6 fold, p = 2.93e-17**	**3, 4.6 fold, p = 0.023**
aut dom	**41, 1.9 fold, p = 0.0002**	**36, 0.7 fold, p = 0.044**	8, 0.9 fold, p = 0.8
aut rec	**46, 0.7 fold, p = 0.008**	**177, 1.2 fold, p = 0.008**	**9, 0.3 fold, p = 2e-06**
x-linked	17, 0.8fold, p = 0.4	43, 0.8 fold, p = 0.3	**30, 3.3 fold, p = 4.3e-11**
	individual classes		
	C1 – 36	**C2 – 60**	**C3 – 10**
hPSD	**8, 3.1 fold, p = 0.001**	**10, 2.3 fold, p = 0.009**	**4, 5.6 fold, p = 0.0006**
HI	**6, 11.4 fold, p = 7.5e-12**	1, 1.1 fold, p = 1	0, 0 fold, p = 1
aut dom	11, 1.5 fold, p = 0.23	9, 0.7 fold, p = 0.4	1, 0.5 fold, p = 0.67
aut rec	16, 0.8 fold, p = 0.15	42, 1.2 fold, p = 0.12	**2, 0.3 fold, p = 0.035**
x-linked	9, 1.2 fold, p = 0.66	13, 1.1 fold, p = 0.95	**7, 3.4 fold, p = 0.0007**
	C4 - 59	**C5 - 112**	**C6 - 35**
hPSD	6, 1.4 fold, p = 0.51	**23, 2.9 fold, p = 1.14e-07**	**9, 3.6 fold, p = 8.12e-05**
HI	**10, 11.6 fold, p = 2.7e-20**	**13, 8 fold, p = 5.1e-17**	**3, 5.9 fold, p = 0.0052**
aut dom	19, 1.6 fold, p = 0.07	18, 0.8 fold, p = 0.4	7, 1 fold, p = 1
aut rec	30, 0.9 fold, p = 0.33	77, 1.2 fold, p = 0.06	**7, 0.3 fold, p = 2.5e-05**
x-linked	10, 0.8 fold, p = 0.66	21, 0.9 fold, p = 0.81	**23, 3.2 fold, p = 6.6e-09**
	C7 - 21	**C8 - 104**	**C9 - 1**
hPSD	1, 0.7fold, p = 1	**19, 2.6fold, p = 2.57e-05**	0, 0fold, p = 1
HI	**4, 13.1fold, p = 7.7e-09**	**9, 5.9fold, p = 1.8e-08**	**1, 68.5fold, p = 5.3e-05**
aut dom	**21, 3.2fold, p = 3.9e-06**	**11, 0.5fold, p = 0.03**	0, 0fold, p = 1
aut rec	**4, 0.3fold, p = 0.0009**	71, 1.2fold, p = 0.09	0, 0fold, p = 0. 9
x-linked	1, 0.2fold, p = 0.14	14, 0.7fold, p = 0.15	1, 4.9fold, p = 0.5

Result of enrichment analyses for the six major groups and the nine individual classes of ID genes against the human genome (~20500 genes). Additional to inheritance modes from the ID catalogue database (in green), a dataset of 1458 hPSD genes[322] (in blue), and a dataset of 299 HI genes[323] (in red) were used; behind the group or class the number of ID genes within the group or class in given; order of results: number of mapped genes, fold enrichment, uncorrected p-value with chi-square test; bold letters indicate results with significant p-values.

No significant enrichment of inheritance patterns was found in the CS, CM, and NC groups.

The SWSM group showed significant enrichment for autosomal dominant inheritance (1.9 fold), while it was depleted for autosomal recessive inheritance (0.7 fold). The SWOSM group was significantly enriched for autosomal recessive inheritance (1.2 fold), while it was significantly depleted for autosomal dominant inheritance (0.7 fold). The non-syndromic group was highly enriched for X-linked inheritance (3.3 fold), while it was significantly depleted for autosomal recessive inheritance (0.3 fold) (table 4).

Regarding individual classes, classes 3 and 6 were significantly enriched for X-linked inheritance (3.5 fold and 3.2 fold respectively), and class 7 was significantly enriched for autosomal dominant inheritance (3.2 fold). Classes 3, 6, and 7 were significantly depleted for autosomal recessive inheritance, and class 8, in particular its subclass 8b, was significantly depleted for autosomal dominant inheritance (table 4).

Expression patterns

When analyzing expression levels of ID genes against the whole genome background, the classic severe group showed significantly enriched expression patterns for different brain regions. The genes from the classic mild to moderate group did not show significantly enriched expression levels, and the non-classic group showed significantly increased expression in several abdominal organs as well as in various blood and immune cells (table 5).

While the SWSM group did not show significantly increased expression levels, genes from the SWOSM group were significantly enriched in blood and immune cells, abdominal organs, and brain regions. The NS group showed significantly increased expression exclusively in several brain regions.

Regarding individual classes, class 2 and class 6 showed significantly enriched expression in various brain regions, while class 1 showed a wide and heterogeneous expression pattern in brain and organs and class 8b in brain regions, organs and blood and immune cells. Class 5 showed significantly increased expression in liver and CD34+ cells.

When comparing the classes against the background of all 388 ID genes, the CS group showed significantly increased expression in fetal brain, whereas CM and NC groups showed no significant expression enrichment. SWSM, SWOSM, and NS groups showed no significant expression pattern, either.

Regarding individual classes, class 1 genes showed significantly higher expression in various brain regions and tissues (fetal brain, dorsal root ganglion, trigeminal ganglion, prefrontal cortex, ciliary ganglion, globus pallidus, skin, cerebellum, atrio ventricular node, cerebellum peduncles, retina, adrenal cortex, superior cervical ganglion, medulla oblongata, colorectal carcinoma, subthalamic nucleus, occipital lobe, cingulate cortex, temporal lobe,

ovary, spinal cord, caudate nucleus, appendix, and olfactory bulb). Class 4 genes showed significantly lower expression in several blood cell lines (BDCA4+ dendritic cells, CD105+ endothelial, X721_B lymphoblasts, CD34+, CD56+ NK cells, CD33+ myeloid, and CD14+ monocytes). Class 8b genes showed significant increased expression in X721_B lymphoblasts, CD33+ myeloid, BDCA4+ dendritic cells, CD14+ monocytes, CD56+ NK cells, CD34+, CD105+ endothelial, CD8+ T-cells, CD4+ T-cells, Cd19+ B-cells (neg. sel.), leukemia promyelocytic HL-60, colon, small intestine, thyroid, lymph node, adrenal gland, heart, lymphoma burkitts (Daudi), and thymus).

Table 5 Significant enrichment of expression in 84 different human tissues and cell types against the background of the human genome (13100 genes)

<table>
<tr><th colspan="3">classic and non-classic groups</th></tr>
<tr><td colspan="3">CS – 102 (92)
fetal brain, amygdala, whole brain, prefrontal cortex, occipital lobe, globus pallidus, medulla oblongata, parietal lobe, temporal lobe, cingulate cortex, subthalamic nucleus, pineal gland (night), cerebellum, caudate nucleus, pineal gland (day), cerebellum peduncles, hypothalamus, thalamus, pons, retina</td></tr>
<tr><td colspan="3">CM – 200 (175)
no significant enrichment of expression in specific tissues</td></tr>
<tr><td colspan="3">NC – 126 (111)
liver, CD33+ myeloid, X721_B lymphoblasts, colon, CD105+ endothelial, CD34+, thyroid, CD14+ monocytes, kidney, pineal gland (night), BDCA4+ dendritic cells, small intestine, CD56+ NK cells, bronchial epithelial cells, pineal gland (day), fetal liver, CD8+ T cells and leukemia promyelocytic-HL-60</td></tr>
<tr><th colspan="3">syndromic and non-syndromic groups</th></tr>
<tr><td>SWSM – 106 (92)
no significant expression</td><td>SWOSM – 256 (225)
liver, CD34+, X721_B lymphoblasts, pineal gland (night), leukemia promyelocytic-HL-60, CD105+ endothelial, leukemia chronic myelogenous K-562, kidney, pineal gland (day), BDCA4+ dendritic cells, whole brain, thyroid, lymphoma burkitts (Daudi), amygdala, fetal liver, leukemia, lymphoblastic (MOLT-4), CD56+ NK cells, hypothalamus, CD33+ myeloid, adipocyte, prefrontal cortex, occipital lobe, bronchial epithelial cells, CD19+ B cells (neg. sel.), caudate nucleus, small intestine, thalamus, CD14+ monocytes, colon, spinal cord</td><td>NS – 45 (38)
fetal brain, prefrontal cortex, amygdala, occipital lobe, cingulate cortex</td></tr>
<tr><th colspan="3">individual classes</th></tr>
<tr><td>C1 – 36 (35)
fetal brain, prefrontal cortex, cerebellum, occipital lobe, globus pallidus, amygdala, cerebellum peduncles, dorsal root ganglion, medulla oblongata, whole brain, subthalamic nucleus, retina, caudate nucleus, temporal lobe, trigeminal ganglion, spinal cord, cingulate cortex, olfactory bulb, hypothalamus, colorectal adenocarcinoma, ciliary ganglion, parietal lobe, colon, pineal gland (night), pineal gland (day), thalamus, atrio-ventricular node, pons, skin, ovary, adrenal cortex, superior cervical ganglion, uterus, adrenal gland, skeletal muscle, adipocyte, uterus corpus, prostate, trachea, appendix,</td><td>C2 – 60 (51)
amygdala, whole brain, occipital lobe</td><td>C3 – 10 (8)
no significant expression</td></tr>
</table>

kidney, small intestine, testis leydig cell, pituary		
C4 – 59 (47) no significant expression	**C5 – 112 (103)** liver, CD34+	**C6 – 35 (30)** prefrontal cortex, fetal brain, cingulate cortex, amygdala, occipital lobe, subthalamic nucleus, medulla oblongata, temporal lobe
C7 – 21 (20) no significant expression	**C8 - 104** 8a (42): no significant expression 8b (53): X721_B-lymphoblasts, CD34+, CD33+ myeloid, BDCA4+ dendritic cells, CD105+ endothelial, leukemia promyelocytic-HL-60, CD56+ K cells, CD14+ monocytes, thyroid, colon, kidney, CD8+ T-cells, liver, lymphoma burkitts (Daudi), adrenal gland, small intestine, CD4+ T-cells, CD19+ B-cells (neg. sel.), fetal liver, pancreatic islet, pineal gland (night), leukemia lymphoblastic (MOLT-4), bronchial epithelial cells, lymphoma burkitts (Raji), heart, adipocyte, pineal gland (day), fetal thyroid, lymph node, leukemia-chronic myelogenous K-562, thymus, whole blood, tongue, prostate, CD71+ early erythroid, hypothalamus, salivary gland, colorectal adenocarcinoma, whole brain, adrenal cortex	**C9 – 1 (0)**

numbers in brackets show number of genes with available expression data; tissues are given in the order of increasing p-values (all given with a p-value <0.05)

DISCUSSION

The need for an ID gene catalogue and a database

To cope with the arising challenges resulting from NGS technologies and to identify and characterize the complex networks of cognitive function and dysfunction, a systematic, large-scale and collaborative interdisciplinary approach is required.

In the past, several attempts have been undertaken to create comprehensive lists of ID genes. In 2003, Inlow and Restifo[257] collected 282 ID associated genes from the OMIM database and through literature search. They also developed a biological functions classification scheme including information on metabolic pathways, signaling pathways, transcription and other aspects of neuronal and glial function. Furthermore, they investigated the conservation of human ID genes in *Drosophila melanogaster.*[257] Betancur[325] and Kou et al.[96] developed curated lists of 114 genes for autism spectrum disorders (ASD) and 223 ID genes, respectively. The ASD gene list is reported to be exhaustive, the overlapping ID gene list to be diverse but not exhaustive. These lists were used to prioritize genes and pathways for ASD and ID.[96]

All of these lists do either contain an incomplete inventory of ID genes or did not fulfill our requirements for reliability of data e.g. via independent evidence. We therefore aimed at an own, systematic catalog containing all genes reliably underlying ID, thus not including genes and disorders with clinically or genetically low evidence, treatable metabolic conditions, clearly neurodegenerative disorders or disorders with early lethality.

This list is integrated into a database, supplemented with various information on associated disorders and inheritance patterns, and connected to various datasets from both human and animal models. After publication the database will be public.

The list and the database are supposed to meet different aims: 1) To provide a repository of genes, mutations in which are sufficient to cause ID. The catalog is supposed to be regularly updated, manually curated, and accessible to others. 2) This gene catalog provides the basis for large-scale functional screens of ID genes in *Drosophila melanogaster* in order to gain more insight into neuronal function and dysfunction. A first approach has examined visual behavior, photoreceptor physiology and multiparametric morphological characteristics, resulting in a high number of novel eye phenotypes and the identification of highly connected functional modules.[326] 3) Data from these screens is supposed to allow mapping of novel functional networks. This might be contributing to recognition of common themes underlying ID and to the prediction of disease genes in short-term, and to the establishment of therapeutic interventions in long-term. 4) The list can be used for diagnostics. Whole exome sequencing usually results in a large number of potential variants. These can be compared to the ID gene catalog in order to facilitate interpretation of results. As the main list does only contain highly reliable ID genes fulfilling certain criteria, for this purpose the secondary list with a larger number of less reliable, but potential ID genes, can be used additionally.

The periodically updated catalog might also be a nice tool to reflect the changes in ID genetics. So far, the majority of ID genes follow an autosomal-recessive inheritance pattern, and also the proportion of X-linked genes is quite high. This is probably due to the fact that gene identification for autosomal-recessive and X-linked ID genes was facilitated by many recessive genes being involved in recognizable metabolic disorders and that familial ID allowed the usage of linkage analyses for quite some time. Autosomal-dominant genes are underrepresented as sporadic ID is usually caused by *de novo* mutations. These could not be systematically searched for until recently. With NGS technologies large-scale screens for *de novo* mutations have now become possible.[34,35,42] It will be interesting to see if these advancing changes in technologies will be reflected in the proportion of inheritance modes of ID genes over time. When comparing the set of 388 ID genes from 2010 with the set of 518 ID genes from May 2013, only mild changes can be seen. The fraction of autosomal recessive genes is unaltered (59%), probably due to the successful combination of linkage

analysis or homozygosity mapping with NGS technologies in consanguineous ID families. The number of mitochondrial genes is unaltered, and the fraction of X-linked genes slightly decreased from 23% to 20%. Related to their small genomic proportion there are probably only few X-linked and mitochondrially encoded genes remaining to be still identified. The impact of NGS on increasing identification of genes with *de novo* mutations in relation to other inheritance modes is not yet very obvious but potentially reflected in a mild increase from 21% to 24% autosomal dominant ID genes. In this regard a continuously increasing proportion is to be expected for the future.

The rationale for a classification

Though a definition of ID exists, there is no clear definition of what constitutes an ID disorder. ID disorders are not only genetically but also clinically extremely heterogeneous. During the last years many examples were found to support the idea of similar phenotypes being caused by defects in functionally related genes.[178-180] We therefore aimed at a phenotypic classification of ID disorders using a manageable and comprehensive amount of phenotype classes based on clinical manifestation and severity. The idea was that these groups, classified according to specific phenotypic features, might mirror disruptions of certain molecular functions or processes.

So far, similar attempts are limited to a description of frequency and nature of symptoms in ID disorders. An example is the Human Phenotype Ontology (HPO)[327,328] which aims to establish a standardized, controlled vocabulary for phenotypic information in order to exploit semantic similarities for database searches for clinical diagnostics or for large-scale computational analysis of gene expression patterns.[327] HPO uses clinical information from OMIM and considers frequencies but is not manually curated. Also OMIM provides an overview of clinical symptoms via the “clinical synopsis” function, but no data on frequencies is available there.

Our approach is the first to use manually curated data and a newly introduced classification system comprising nine individual classes and six higher-order overlapping groups according to severity and nature of ID on one hand and according to syndromic or non-syndromic occurrence on the other hand (figure 1). From this sorting we expect more reliable computational analyses compared to using a large amount of more detailed, but not manually curated, error-prone data.

Of course, the classification is highly subjective and depending on the respective clinician. To minimize this “human” effect and to provide consistency, all genes were classified by the same clinician, the author of this thesis, and all entries were revised by a second, independent clinician. In addition, we also used information on symptoms with frequencies being taken into account (table 1). Also here, highly reliable, manually curated

data is used. Another possible objection, which however would apply to all kind of such classifications, is the availability and detailedness of the data underlying the classification. Depending on the background of the authors and the aim of the respective reports in literature and databases, biases might exist, and clinical information can extremely vary in amount and quality. To take into account limitations in clinical data for later enrichment analyses, we marked those entries accordingly (letter Z in additional phenotype information (table 1)).

First enrichment patterns

Regarding the first analyzed datasets, several patterns could already be deduced.

In general, ID genes are enriched for hPSD genes and HI genes when compared to the whole genome background. Both observations were to be expected as it has been shown previously that a high number of brain diseases involves hPSD genes[322] and as many of the phenotypes associated with haploinsufficiency include ID and developmental disorders.[323]

Of note, patterns for haploinsufficiency, localization in the postsynaptic density, and inheritance could rather be found for the syndromic versus non-syndromic ID groups than for the the classic versus non-classic groups. Interesting patterns regarding expression data could be found for both classic/non-classic groups as well as for syndromic/non-syndromic groups. These first observations of patterns between the six higher- order groups confirm the rationale for the classification.

Some major patterns were identified in the current dataset:

1) **hPSD genes are highly significantly enriched in the non-syndromic group (NS), while they are not enriched in the SWSM group.** This is in accordance with significantly increased expression levels for NS genes in fetal brain and several brain regions. The brain specific localization and function of NS genes goes in hand with rather brain specific, non-syndromic manifestation of ID. In contrast, SWSM genes did not show significantly enriched expression patterns, and the SWOSM genes rather show a heterogeneous expression pattern.

   Of note, X-linked inheritance is enriched in the NS group, whereas autosomal recessive inheritance is significantly depleted. This might indicate a high proportion of synaptic genes on the X-chromosome. This was indeed discussed in a previous study that found 28% of X-linked ID genes to encode postsynaptic genes and 49% of X-chromosomal PSD genes to be involved in human psychiatric disorders.[329] Our data with 20% of X-linked genes in the NS group being hPSD genes and with 46% of hPSD genes in the NS group being X-linked genes, confirms this observation and additionally points to non-syndromic ID as the associated clinical manifestation.

Whereas also numerous other reports noted an increasing role of genes with synaptic function in ID,[3,4,49,77,78,317] the significance of this observation has remained controversial.[63] Here we show that the frequency of synaptic genes amongst ID genes might correlate with the associated clinical phenotype and/or the inheritance mode. Diversity of clinical phenotypes and inheritance modes between different studies due to specific selection criteria for the patient cohorts might therefore explain discrepancies in their observations and further confirms the value of our classification system.

2) **The highest enrichment of HI genes is found in the SWSM group.** This is in accordance with the significant enrichment of autosomal dominant inheritance (in terms of *de novo* mutations) in the SWSM group, while it is underrepresented or not enriched in the SWOSM and NS groups. This observation emphasizes the role of HI genes in disorders that affect brain and organ development.

   Of note, the SWSM group does not show any significant expression pattern, though specific organ or other tissue expression might have been expected due to the broad syndromic phenotypic spectrum. However, when looking at the individual classes, class 1 (SWSM) indeed shows significantly high expression for a large number of brain and other tissues, while class 2 (SWOSM) and class 6 (NS), that are not associated with structural malformations, rather show brain specific expression.

3) **Autosomal-recessive inheritance is significantly enriched in the SWOSM group while it is significantly depleted in the SWSM and the NS groups.** This indicates that mutations in autosomal recessive genes predominantly cause syndromic ID but without structural malformations. This is true for most metabolic disorders, which are indeed predominantly in the SWOSM group (data not shown). The expression pattern of SWOSM genes is very heterogeneous. Significant expression of genes in class 8b which contains many mitochondrial disorders (data not shown) includes several blood or tumor cells as well as organs that are typically affected in mitochondrial disorders: liver, heart, and brain.

4) **Expression in brain correlates with severity of ID.** CS genes show increased expression in various brain regions, while CM genes show no significant expression pattern, and NC genes are heterogeneously expressed in organs, brain tissues, and blood or immune cells. Thus, the high and specific expression of CS genes in brain is correlating with the severity of ID.

Conclusion and outlook

The above described analyses and observations are only one part and the first step to predict disease genes and molecular networks in a systems biology-like approach. While it is

not the first attempt to systematically pool and use data on ID genes, it is the first approach using highly reliable, manually curated data, a novel clinical classification system and merging it with both publically available datasets and with systematic functional data from model organism screens (e.g. Oortveld et al.,[326]).
First enrichment analyses in this study already show patterns that might be of help to achieve these goals in the near future.

Supplementary table

gene	inheritance	inheritance	main class	additional phenotypes	disorder
ABCD1	Mendelian X-linked	not sure	8a, 8b	L, G, K, H, M, P	ADRENOLEUKODYSTROPHY
ABHD5	Mendelian autosomal	recessive	8a	M, Q, S	CHANARIN-DORFMAN SYNDROME
ACOX1	Mendelian autosomal	recessive	8b	M, H, E, G, C, L	Peroxisomal acyl-CoA oxidase deficiency
ACSL4	Mendelian X-linked	recessive	6		MENTAL RETARDATION, X-LINKED
ACVR1	Mendelian autosomal	dominant	7	A, Ub	FIBRODYSPLASIA OSSIFICANS PROGRESSIVA
ADCK3	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	Primary coenzyme Q10 deficiency
ADSL	Mendelian autosomal	recessive	5	M, P, E	Adenylosuccinase deficiency
AFF2	Mendelian X-linked	not sure	6	P	MENTAL RETARDATION, X-LINKED
AGA	Mendelian autosomal	recessive	8b	M, G, H, (S), (C)	ASPARTYLGLUCOSAMINURIA
AGPAT2	Mendelian autosomal	recessive	8a	M, Q, K	LIPODYSTROPHY, CONGENITAL GENERALIZED
AGTR2	Mendelian X-linked	recessive	3	E, (P)	Mental retardation, X-linked
AHCY	Mendelian autosomal	recessive	5	M, H, (Q), (L)	HYPERMETHIONINEMIA WITH S-ADENOSYLHOMOCYSTEINE HYDROLASE DEFICIENCY
AHI1	Mendelian autosomal	recessive	4	H, L, O, T, W	Joubert syndrome
AK1	Mendelian autosomal	recessive	8a	M, R	Hemolytic anemia due to adenylate kinase deficiency
ALDH18A1	Mendelian autosomal	recessive	5	A, H, S, T, (M)	CUTIS LAXA
ALDH3A2	Mendelian autosomal	recessive	2	H, S, T, M, A, E	SJOGREN-LARSSON SYNDROME
ALDH5A1	Mendelian autosomal	recessive	5	M, E, H, L, P	SUCCINIC SEMIALDEHYDE DEHYDROGENASE DEFICIENCY
ALG1	Mendelian autosomal	recessive	2	M, E, B, R, L, J, C, W, (Q)	Congenital disorder of glycosylation
ALG12	Mendelian autosomal	recessive	2	M, H, B, R, J, E, V	Congenital disorder of glycosylation
ALG2	Mendelian autosomal	recessive	2	M, E, T	Congenital disorder of glycosylation
ALG3	Mendelian autosomal	recessive	2	M, E, B, T, L	Congenital disorder of glycosylation
ALG6	Mendelian autosomal	recessive	5	M, E, H, T, K	Congenital disorder of glycosylation
ALG9	Mendelian autosomal	recessive	5	M, H, E, B	Congenital disorder of glycosylation
AMT	Mendelian autosomal	recessive	8b	M, H, E, C, (G)	GLYCINE ENCEPHALOPATHY
ANO10	Mendelian autosomal	recessive	8a	G, L, N, Q	SPINOCEREBELLAR ATAXIA, AUTOSOMAL RECESSIVE
AP1S1	Mendelian autosomal	recessive	2	(M), H, C, S, O	MENTAL RETARDATION, ENTEROPATHY, DEAFNESS, PERIPHERAL NEUROPATHY, ICHTHYOSIS, AND KERATODERMA
AP1S2	Mendelian X-linked	recessive	6	P, (L)	MENTAL RETARDATION, X-LINKED, SYNDROMIC
AP3B1	Mendelian autosomal	recessive	8a	T, R, J	Hermansky-Pudlak syndrome
APTX	Mendelian autosomal	recessive	5	H, G, L	ATAXIA, EARLY-ONSET, WITH OCULOMOTOR APRAXIA AND HYPOALBUMINEMIA
ARFGEF2	Mendelian autosomal	recessive	2	B, L, (H), (E), (J)	Periventricular heterotopia
ARHGEF6	Mendelian X-linked	recessive	6		MENTAL RETARDATION, X-LINKED
ARL13B	Mendelian autosomal	recessive	4	L, (T), (O)	JOUBERT SYNDROME
ARL6	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
ARX	Mendelian X-linked	recessive	1, 2, 3, 5	E, H, L	MENTAL RETARDATION, X-LINKED
ARX	Mendelian X-linked	recessive	1, 2, 3, 5	E, H, L	EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE
ARX	Mendelian X-linked	recessive	1, 2, 3, 5	E, H, L	PARTINGTON X-LINKED MENTAL RETARDATION SYNDROME
ARX	Mendelian X-linked	recessive	1, 2, 3, 5	E, H, L	CORPUS CALLOSUM, AGENESIS OF, WITH ABNORMAL GENITALIA
ARX	Mendelian X-linked	recessive	1, 2, 3, 5	E, H, L, C	Lissencephaly
ASL	Mendelian autosomal	recessive	8a	M, H, P, E, S	Urea Cycle Disorders, ARGININOSUCCINIC ACIDURIA
ASPA	Mendelian autosomal	recessive	2, 5	M, F, H, G, C, L	CANAVAN DISEASE
ASPM	Mendelian autosomal	recessive	5	BBc, (L)	MICROCEPHALY, PRIMARY, AUTOSMAL RECESSIVE
ATP1A2	Mendelian autosomal	dominant	5	H, E, G	ALTERNATING HEMIPLEGIA OF CHILDHOOD
ATP2A2	Mendelian autosomal	dominant	8a	S	DARIER-WHITE DISEASE
ATP6AP2	Mendelian X-linked	recessive	6	E, (H)	MENTAL RETARDATION, X-LINKED, SYNDROMIC

ATP6V0A2	Mendelian autosomal	recessive	1	E, G, H, L, S, T	CUTIS LAXA, AUTOSOMAL RECESSIVE
ATP6V0A2	Mendelian autosomal	recessive	5	E, G, H, L, S, T	WRINKLY SKIN SYNDROME
ATP7A	Mendelian X-linked	recessive	8b	M, E, H, G, C, S	Menkes disease
ATP7A	Mendelian X-linked	recessive	5	S	OCCIPITAL HORN SYNDROME
ATR	Mendelian autosomal	recessive	5	AA, BB	Seckel syndrome
ATRX	Mendelian X-linked	recessive	2	H, U, W, K	MENTAL RETARDATION-HYPOTONIC FACIES SYNDROME, X-LINKED
ATRX	Mendelian X-linked	recessive	2	A, B, W, R, (E)	ALPHA-THALASSEMIA/MENTAL RETARDATION SYNDROME, X-LINKED
AUH	Mendelian autosomal	recessive	5	M, H, G	3-METHYLGLUTACONIC ACIDURIA
B3GALTL	Mendelian autosomal	recessive	4	A, T, Uc, V	PETERS-PLUS SYNDROME
B4GALT1	Mendelian autosomal	recessive	4	M, L, R	Congenital disorder of glycosylation
B4GALT7	Mendelian autosomal	recessive	5	A, S	EHLERS-DANLOS SYNDROME, PROGEROID FORM
BBS1	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BBS10	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BBS12	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BBS2	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BBS4	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BBS5	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BBS7	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BBS9	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
BCKDHA	Mendelian autosomal	recessive	8a	M, (H) (untreated), (C) (untreated)	MAPLE SYRUP URINE DISEASE
BCKDHB	Mendelian autosomal	recessive	8a	M, (H) (untreated), (C) (untreated)	MAPLE SYRUP URINE DISEASE
BCOR	Mendelian X-linked	recessive	4	T, Uab, W, B	MICROPHTHALMIA, SYNDROMIC
BCS1L	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	mitochondrial complex III deficiency
BCS1L	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	LEIGH SYNDROME
BLM	Mendelian autosomal	recessive	5	A, S, I, J, K, C	BLOOM SYNDROME
BRAF	Mendelian autosomal	dominant	4	A, S, V, E, Q, (I)	CARDIOFACIOCUTANEOUS SYNDROME
BRAF	Mendelian autosomal	dominant	4	A, S, V, E, Q, (I)	Leopard syndrome
BRAF	Mendelian autosomal	dominant	4	A, S, V, E, Q, (I)	Noonan syndrome
BRWD3	Mendelian X-linked	recessive	5	F head	MENTAL RETARDATION, X-LINKED
BSCL2	Mendelian autosomal	recessive	5	M, K, Q	LIPODYSTROPHY, CONGENITAL GENERALIZED
BUB1B	Mendelian autosomal	recessive	4	A, B, I, U, V, W, X	MOSAIC VARIEGATED ANEUPLOIDY SYNDROME
CA2	Mendelian autosomal	recessive	5	M, A, W, U	OSTEOPETROSIS, AUTOSOMAL RECESSIVE
CACNA1C	Mendelian autosomal	dominant	4	H, C, V, Ua, J, P	TIMOTHY SYNDROME
CASK	Mendelian X-linked	not sure	1	B, L, H, (A)	MENTAL RETARDATION AND MICROCEPHALY WITH PONTINE AND CEREBELLAR HYPOPLASIA
CASK	Mendelian X-linked	recessive	2	P	FG syndrome
CBS	Mendelian autosomal	recessive	5	M, T, R, F, Ub, E	HOMOCYSTINURIA DUE TO CYSTATHIONINE BETA-SYNTHASE DEFICIENCY
CC2D1A	Mendelian autosomal	recessive	3		MENTAL RETARDATION, AUTOSOMAL RECESSIVE
CC2D2A	Mendelian autosomal	recessive	4	T, L, O	Joubert syndrome
CC2D2A	Mendelian autosomal	recessive	4	T, L, O	COACH SYNDROME
CDK5RAP2	Mendelian autosomal	recessive	5	BBc	MICROCEPHALY, PRIMARY, AUTOSOMAL RECESSIVE
CDKL5	Mendelian X-linked	not sure	2	E, B, H	ANGELMAN SYNDROME
CDKL5	Mendelian X-linked	not sure	2	E, H, Ba, O, P	EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE
CENPJ	Mendelian autosomal	recessive	8a	AAc, BBc, Z	Seckel syndrome
CENPJ	Mendelian autosomal	recessive	2	BBc, (E)	MICROCEPHALY, PRIMARY, AUTOSOMAL RECESSIVE
CEP290	Mendelian autosomal	recessive	4	L, H, O, T, W, (H), (Ua)	Joubert syndrome
CEP290	Mendelian autosomal	recessive	4	L, H, O, T, W, (H), (Ua)	Bardet-Biedl syndrome
CHD7	Mendelian autosomal	dominant	4	A, V, T, W, X, Uc	CHARGE SYNDROME
CNTNAP2	Mendelian autosomal	recessive	2	E, O, G, P	CORTICAL DYSPLASIA-FOCAL EPILEPSY SYNDROME

CNTNAP2	Mendelian autosomal	dominant	6		AUTISM, SUSCEPTIBILITY TO
COG1	Mendelian autosomal	recessive	5	M, A, B, L, V	Congenital disorder of glycosylation
COG7	Mendelian autosomal	recessive	8b	M, H, C, E, S, U	Congenital disorder of glycosylation
COG8	Mendelian autosomal	recessive	2	B, E, H, M	Congenital disorder of glycosylation
COL4A1	Mendelian autosomal	dominant	7	H, E, L, T	PORENCEPHALY
COQ2	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	COENZYME Q10 DEFICIENCY, PRIMARY
COX10	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	Encephalopathy, progressive mitochondrial, with proximal renal tubulopathy due to cytochrome c oxidase deficiency
COX15	Mendelian autosomal	recessive	8b	M, L, H, G, Q	LEIGH SYNDROME
COX15	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	CARDIOENCEPHALOMYOPATHY, FATAL INFANTILE, DUE TO CYTOCHROME c OXIDASE DEFICIENCY
CRBN	Mendelian autosomal	recessive	6		MENTAL RETARDATION, AUTOSOMAL RECESSIVE
CREBBP	Mendelian autosomal	dominant	1	A, T, V, W, Ua	RUBINSTEIN-TAYBI SYNDROME
CTDP1	Mendelian autosomal	recessive	5	A, H, T, K, Q, L, U, G	CONGENITAL CATARACTS, FACIAL DYSMORPHISM, AND NEUROPATHY
CUL4B	Mendelian X-linked	recessive	5	A, H, K, N, P	MENTAL RETARDATION, X-LINKED, WITH SHORT STATURE, HYPOGONADISM, AND ABNORMAL GAIT
CYB5R3	Mendelian autosomal	recessive	2	M, A, B, G, H, C, E	METHEMOGLOBINEMIA DUE TO DEFICIENCY OF METHEMOGLOBIN REDUCTASE
D2HGDH	Mendelian autosomal	recessive	5	M, E, H, (L), (Q)	D-2-HYDROXYGLUTARIC ACIDURIA
DARS2	Mendelian autosomal	recessive	8a	L, M, H, G	LEUKOENCEPHALOPATHY WITH BRAINSTEM AND SPINAL CORD INVOLVEMENT AND LACTATE ELEVATION
DBT	Mendelian autosomal	recessive	8a	M, H, (C) (if untreated)	MAPLE SYRUP URINE DISEASE
DCX	Mendelian X-linked	not sure	1, 4	B, E, L, P, C	LISSENCEPHALY
DHCR24	Mendelian autosomal	recessive	1	M, V, Uc, A	DESMOSTEROLOSIS
DHCR7	Mendelian autosomal	recessive	1	M, A, B, Ua, Uc, V, W, P	SMITH-LEMLI-OPITZ SYNDROME
DIP2B	Mendelian autosomal	dominant	5, 6		MENTAL RETARDATION
DKC1	Mendelian X-linked	recessive	4	A, B, I, J, R, H, L, S	HOYERAAL-HREIDARSSON SYNDROME
DKC1	Mendelian X-linked	recessive	8a	I, J, R, H, S	DYSKERATOSIS CONGENITA
DLD	Mendelian autosomal	recessive	5	M, H, (C) (if untreated)	DIHYDROLIPOAMIDE DEHYDROGENASE DEFICIENCY
DLG3	Mendelian X-linked	recessive	3		MENTAL RETARDATION, X-LINKED
DMD	Mendelian X-linked	recessive	8a	C, Q	MUSCULAR DYSTROPHY
DMPK	Mendelian autosomal	dominant	8a	Q, T, (C)	MYOTONIC DYSTROPHY
DNAJC19	Mendelian autosomal	recessive	5	A, M, H, Q, W, R	3-METHYLGLUTACONIC ACIDURIA
DNMT3B	Mendelian autosomal	recessive	5	J	IMMUNODEFICIENCY-CENTROMERIC INSTABILITY-FACIAL ANOMALIES SYNDROME
DPAGT1	Mendelian autosomal	recessive	5	M, H, E, B	Congenital disorder of glycosylation
DPM1	Mendelian autosomal	recessive	5	M, B, E, H	Congenital disorder of glycosylation
DPYD	Mendelian autosomal	recessive	5	M, E, H	DIHYDROPYRIMIDINE DEHYDROGENASE DEFICIENCY
DYM	Mendelian autosomal	recessive	5	AA, B, U	DYGGVE-MELCHIOR-CLAUSEN DISEASE
EHMT1	Mendelian autosomal	dominant	1	H, V, W, J, E, P, N, B	KLEEFSTRA SYNDROME
EIF2AK3	Mendelian autosomal	recessive	8a	AA, C, K, U	EPIPHYSEAL DYSPLASIA, MULTIPLE, WITH EARLY-ONSET DIABETES MELLITUS
EMX2	Mendelian autosomal	dominant	4	L, H, E	SCHIZENCEPHALY
EP300	Mendelian autosomal	dominant	1	A, (N), (Ua), (T), (V), (W)	RUBINSTEIN-TAYBI SYNDROME
ERCC2	Mendelian autosomal	recessive	8b	B, H, G, S, T, I	XERODERMA PIGMENTOSUM
ERCC2	Mendelian autosomal	recessive	2	A, B, T, W, U, S, C	CEREBROOCULOFACIOSKELETAL SYNDROME
ERCC2	Mendelian autosomal	recessive	5	B, H, G, S, T, I	TRICHOTHIODYSTROPHY
ERCC3	Mendelian autosomal	recessive	8a	B, H, G, S, T, I	XERODERMA PIGMENTOSUM
ERCC3	Mendelian autosomal	recessive	5	B, H, G, S, T, I	TRICHOTHIODYSTROPHY
ERCC5	Mendelian autosomal	recessive	8b	B, H, G, S, T, I	XERODERMA PIGMENTOSUM

ERCC6	Mendelian autosomal	recessive	5	Aa, L, H, G, S, T, C	COCKAYNE SYNDROME
ERCC6	Mendelian autosomal	recessive	5	Aa, L, H, G, S, T, C	DE SANCTIS-CACCHIONE SYNDROME
ERCC6	Mendelian autosomal	recessive	8b	Aa, L, H, G, S, T, C	CEREBROOCULOFACIOSKELETAL SYNDROME
ERCC8	Mendelian autosomal	recessive	5	Aa, L, H, G, S, T, C	COCKAYNE SYNDROME
ESCO2	Mendelian autosomal	recessive	4	Ac, C, Ua, Uc	ROBERTS SYNDROME
ESCO2	Mendelian autosomal	recessive	4	Ac, C, Ua, Uc	SC PHOCOMELIA SYNDROME
ETHE1	Mendelian autosomal	recessive	2, 8b	M, C, G, H, (S), (O)	ENCEPHALOPATHY, ETHYLMALONIC
FANCD2	Mendelian autosomal	recessive	8a	R, A, Ua, I, C, P	FANCONI ANEMIA
FGD1	Mendelian X-linked	not sure	4	A, W, P	AARSKOG-SCOTT SYNDROME
FGFR2	Mendelian autosomal	dominant	4	Ua, Ub	APERT SYNDROME
FGFR2	Mendelian autosomal	dominant	4	Ua, Ub, S, W	BEARE-STEVENSON CUTIS GYRATA SYNDROME
FGFR2	Mendelian autosomal	dominant	7	Ua, Ub, X, (E, C)	PFEIFFER SYNDROME
FGFR2	Mendelian autosomal	dominant	7	Ua, Ub, T	SAETHRE-CHOTZEN SYNDROME
FGFR2	Mendelian autosomal	dominant	4	Ua	SCAPHOCEPHALY, MAXILLARY RETRUSION, AND MENTAL RETARDATION
FGFR3	Mendelian autosomal	dominant	7	Ua, Ub	MUENKE SYNDROME
FH	Mendelian autosomal	recessive	2	M, H, G, C, L, P, E, B	FUMARASE DEFICIENCY
FKRP	Mendelian autosomal	recessive	1	Q, L, T, (C)	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH OR WITHOUT MENTAL RETARDATION)
FKRP	Mendelian autosomal	recessive	1	Q, L, T, (C)	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH BRAIN AND EYE ANOMALIES)
FKTN	Mendelian autosomal	recessive	4	Q, L, H, C, T	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH BRAIN AND EYE ANOMALIES)
FLNA	Mendelian X-linked	dominant	5	L, E	HETEROTOPIA, PERIVENTRICULAR
FLNA	Mendelian X-linked	recessive	5	L, E	FG SYNDROME
FLNA	Mendelian X-linked	not sure	1	Ua, Ub, X, W, V, C, (L)	OTOPALATODIGITAL SYNDROME
FMR1	Mendelian X-linked	not sure	2, 3	P	FRAGILE X MENTAL RETARDATION SYNDROME
FOXG1	Mendelian autosomal	dominant	2	Ba, E, P, L	RETT SYNDROME, CONGENITAL VARIANT
FOXP1	Mendelian autosomal	dominant	6	P	MENTAL RETARDATION WITH LANGUAGE IMPAIRMENT AND AUTISTIC FEATURES
FRAS1	Mendelian autosomal	recessive	7	T, Ua, W, X	FRASER SYNDROME
FTO	Mendelian autosomal	recessive	1	AAc, B, C, L, E, V, W	GROWTH RETARDATION, DEVELOPMENTAL DELAY, COARSE FACIES, AND EARLY DEATH
FTSJ1	Mendelian X-linked	recessive	3	P	MENTAL RETARDATION, X-LINKED
FUCA1	Mendelian autosomal	recessive	8b	M, A, J, H, U, S, E, G, (C)	FUCOSIDOSIS
GAD1	Mendelian autosomal	recessive	2	H	CEREBRAL PALSY, SPASTIC QUADRIPLEGIC
GALE	Mendelian autosomal	recessive	8a	M	GALACTOSE EPIMERASE DEFICIENCY
GALT	Mendelian autosomal	recessive	8b	M, H, (G) without treatment	GALACTOSEMIA
GAMT	Mendelian autosomal	recessive	5	M, E, P, H	GUANIDINOACETATE METHYLTRANSFERASE DEFICIENCY
GATM	Mendelian autosomal	recessive	5	M, P, (E)	ARGININE:GLYCINE AMIDINOTRANSFERASE DEFICIENCY
GCH1	Mendelian autosomal	recessive	8b	M, E, H, G	HYPERPHENYLALANINEMIA
GCSH	Mendelian autosomal	recessive	2	M, H, E, C, (G)	GLYCINE ENCEPHALOPATHY
GDI1	Mendelian X-linked	recessive	6		MENTAL RETARDATION, X-LINKED
GFAP	Mendelian autosomal	dominant	5	L, C, H, E, G, F	ALEXANDER DISEASE
GJC2	Mendelian autosomal	recessive	5	H, L, G	LEUKODYSTROPHY
GJC2	Mendelian autosomal	recessive	5	H, L, G	SPASTIC PARAPLEGIA, AUTOSOMAL RECESSIVE
GK	Mendelian X-linked	recessive	5	M	HYPERGLYCEROLEMIA
GLDC	Mendelian autosomal	recessive	2	M, H, E, C, (G)	GLYCINE ENCEPHALOPATHY

GLI2	Mendelian autosomal	dominant	7	L, Ua, Uc, K, (C)	HOLOPROSENCEPHALY
GLI3	Mendelian autosomal	dominant	7	F, Ua	GREIG CEPHALOPOLYSYNDACTYLY SYNDROME
GLI3	Mendelian autosomal	dominant	7	C, K, Ua, L	PALLISTER-HALL SYNDROME
GNAS	Mendelian autosomal	dominant	5	A, K	PSEUDOHYPOPARATHYROIDISM
GNAS	Mendelian autosomal	dominant	8a	A, K	PSEUDOHYPOPARATHYROIDISM
GNAS	Mendelian autosomal	dominant	8a	A, K, N	PSEUDOPSEUDOHYPOPARATHYROIDISM
GNPAT	Mendelian autosomal	recessive	5	A, M, U, (T)	RHIZOMELIC CHONDRODYSPLASIA PUNCTATA
GNS	Mendelian autosomal	recessive	5, 8b	M, G, P, H	MUCOPOLYSACCHARIDOSIS
GPC3	Mendelian X-linked	not sure	4	F, I, Uabc, V, W, X	SIMPSON-GOLABI-BEHMEL SYNDROME
GPHN	Mendelian autosomal	recessive	8b	E, G, C	MOLYBDENUM COFACTOR DEFICIENCY
GPR56	Mendelian autosomal	recessive	2	E, H, L	POLYMICROGYRIA
GRIA3	Mendelian X-linked	recessive	6	H	MENTAL RETARDATION, X-LINKED
GRIK2	Mendelian autosomal	recessive	6		MENTAL RETARDATION, AUTOSOMAL RECESSIVE
GSS	Mendelian autosomal	recessive	5	M, R, H	GLUTATHIONE SYNTHETASE DEFICIENCY
GTF2H5	Mendelian autosomal	recessive	5	A, S, (H)	TRICHOTHIODYSTROPHY
GUSB	Mendelian autosomal	recessive	5	M, A, U	MUCOPOLYSACCHARIDOSIS
HCCS	Mendelian X-linked	dominant	7	S, T	MICROPHTHALMIA, SYNDROMIC
HESX1	Mendelian autosomal	recessive	7	K, L	SEPTOOPTIC DYSPLASIA
HLCS	Mendelian autosomal	recessive	8b	M, C	HOLOCARBOXYLASE SYNTHETASE DEFICIENCY
HOXA1	Mendelian autosomal	recessive	5	H, O, V	ATHABASKAN BRAINSTEM DYSGENESIS SYNDROME
HPD	Mendelian autosomal	recessive	5	M, H	TYROSINEMIA
HPRT1	Mendelian X-linked	recessive	2	M, H, P	LESCH-NYHAN SYNDROME
HRAS	Mendelian autosomal	dominant	1	A, I, S, V	COSTELLO SYNDROME
HRAS	somatic	N/A	7	T, L, U, S	SCHIMMELPENNING-FEUERSTEIN-MIMS SYNDROME
HSD17B10	Mendelian X-linked	recessive	5	H, P, M	MENTAL RETARDATION, X-LINKED
HSD17B10	Mendelian X-linked	not sure	8b	M, G, H, C	17-BETA-HYDROXYSTEROID DEHYDROGENASE X DEFICIENCY
HSD17B10	Mendelian X-linked	not sure			MENTAL RETARDATION, X-LINKED
IDS	Mendelian X-linked	recessive	8b	M, A, G, C, Ub, F, V, P	MUCOPOLYSACCHARIDOSIS
IDUA	Mendelian autosomal	recessive	5, 8b	M, A, G, C, Ub, T, V	MUCOPOLYSACCHARIDOSIS
IDUA	Mendelian autosomal	recessive	5, 8b	M, A, G, C, Ub, T, V	MUCOPOLYSACCHARIDOSIS
IDUA	Mendelian autosomal	recessive	5, 8a, 8b	M, A, G, C, Ub, T, V	MUCOPOLYSACCHARIDOSIS
IGF1	Mendelian autosomal	recessive	5	AA, K	INSULIN-LIKE GROWTH FACTOR I DEFICIENCY
IKBKG	Mendelian X-linked	dominant	8a	S, T	INCONTINENTIA PIGMENTI
IL1RAPL1	Mendelian X-linked	not sure	6	(P)	MENTAL RETARDATION, X-LINKED
INPP5E	Mendelian autosomal	recessive	5	N, T, W	MORM, MENTAL RETARDATION, TRUNCAL OBESITY, RETINAL DYSTROPHY, AND MICROPENIS
INPP5E	Mendelian autosomal	recessive	5	T, H, L, (W)	JOUBERT SYNDROME
IQSEQ2	Mendelian X-linked	not sure	3	(P), (E)	MENTAL RETARDATION, X-LINKED
KCNJ11	Mendelian autosomal	dominant	5	M, H, E	DIABETES MELLITUS, PERMANENT NEONATAL
KCNK9	Mendelian autosomal	dominant	2	P, H, Q	BIRK-BAREL MENTAL RETARDATION DYSMORPHISM SYNDROME
KDM5C	Mendelian X-linked	recessive	5, 6	A, H, E, P, (B)	MENTAL RETARDATION, X-LINKED
KIAA1279	Mendelian autosomal	recessive	1	B, X, L	GOLDBERG-SHPRINTZEN MEGACOLON SYNDROME
KRAS	Mendelian autosomal	dominant	4	A, V, R, T, S, Q	NOONAN SYNDROME
KRAS	Mendelian autosomal	dominant	4	A, V, R, T, S, Q	CARDIOFACIOCUTANEOUS SYNDROME
KRBOX4	Mendelian X-linked	recessive	6		CHROMOSOME Xp11.3 DELETION SYNDROME
L1CAM	Mendelian X-linked	recessive	1	L, H	HYDROCEPHALUS DUE TO CONGENITAL STENOSIS OF AQUEDUCT OF SYLVIUS
L1CAM	Mendelian X-linked	recessive	4	L, H	MASA SYNDROME
L1CAM	Mendelian X-linked	recessive	4	L, H	CORPUS CALLOSUM, PARTIAL AGENESIS OF, X-LINKED

L2HGDH	Mendelian autosomal	recessive	8b	M, H, P, E, L, F	L-2-HYDROXYGLUTARIC ACIDURIA
LAMA2	Mendelian autosomal	recessive	8a	Q, (E, L)	MUSCULAR DYSTROPHY, CONGENITAL
LAMP2	Mendelian X-linked	not sure	5	M, Q, G, C	DANON DISEASE
LARGE	Mendelian autosomal	recessive	1	Q, L, M, T, C	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH MENTAL RETARDATION)
LARGE	Mendelian autosomal	recessive	1	Q, L, M, T, C	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH BRAIN AND EYE ANOMALIES)
LIG4	Mendelian autosomal	recessive	5	A, B, J, I, S	LIG4 SYNDROME
LRPPRC	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	LEIGH SYNDROME
MAGT1	Mendelian X-linked	not sure	6		MENTAL RETARDATION, X-LINKED
MAGT1	Mendelian X-linked	not sure	5	M	Congenital Disorders of Glycosylation
MAN2B1	Mendelian autosomal	recessive	5	M, G, Q, H, J, L, Ub, C	MANNOSIDOSIS, ALPHA
MANBA	Mendelian autosomal	recessive	5	M	MANNOSIDASE, BETA
MAOA	Mendelian X-linked	recessive	5	P, M	BRUNNER SYNDROME
MAP2K1	Mendelian autosomal	dominant	4	A, S, V	CARDIOFACIOCUTANEOUS SYNDROME
MAP2K2	Mendelian autosomal	dominant	4	A, S, V	CARDIOFACIOCUTANEOUS SYNDROME
MAT1A	Mendelian autosomal	recessive	8a	M	METHIONINE ADENOSYLTRANSFERASE DEFICIENCY
MBD5	Mendelian autosomal	dominant	6	P, (E)	MENTAL RETARDATION, AUTOSOMAL DOMINANT
MCCC1	Mendelian autosomal	recessive	5	M, H, G, E	3-METHYLCROTONYL-CoA CARBOXYLASE DEFICIENCY
MCCC2	Mendelian autosomal	recessive	5, 8b	M, (G)	3-METHYLCROTONYL-CoA CARBOXYLASE DEFICIENCY
MCOLN1	Mendelian autosomal	recessive	2	M, H, T, L, (G)	MUCOLIPIDOSIS
MCPH1	Mendelian autosomal	recessive	5	B, A	MICROCEPHALY, PRIMARY, AUTOSOMAL RECESSIVE
MECP2	Mendelian X-linked	dominant	2, 8b	B, E, G, P, H, A, O, U	RETT SYNDROME
MECP2	Mendelian X-linked	not sure	2, 8b	B, E, G, P, H, A, O, U	ENCEPHALOPATHY, NEONATAL SEVERE
MECP2	Mendelian X-linked	recessive	2, 8b	B, E, G, P, H, A, O, U	LUBS X-LINKED MENTAL RETARDATION SYNDROME
MECP2	Mendelian X-linked	recessive	5, 6	H, P	MENTAL RETARDATION, X-LINKED
MECP2	Mendelian X-linked	recessive	2	B, E, H	ANGELMAN SYNDROME
MED12	Mendelian X-linked	recessive	4	H, L, E, P, V	OPITZ-KAVEGGIA SYNDROME
MED12	Mendelian X-linked	recessive	5	H, P	LUJAN-FRYNS SYNDROME
MEF2C	Mendelian autosomal	dominant	2	E	MENTAL RETARDATION, AUTOSOMAL DOMINANT
MGAT2	Mendelian autosomal	recessive	5	M, E, R, P	CONGENITAL DISORDER OF GLYCOSYLATION
MID1	Mendelian X-linked	recessive	4	L, W, Uc, V, X	OPITZ GBBB SYNDROME
MKKS	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
MKKS	Mendelian autosomal	recessive	7	Ua, W, X, C	MCKUSICK-KAUFMAN SYNDROME
MLL2	Mendelian autosomal	dominant	4	A, H, Uc, V, W, J, (E), (K)	KABUKI SYNDROME
MLYCD	Mendelian autosomal	recessive	5	M, E, H	MALONYL-CoA DECARBOXYLASE DEFICIENCY
MMAA	Mendelian autosomal	recessive	8a, 8b	M, H, W, A	METHYLMALONIC ACIDURIA
MMACHC	Mendelian autosomal	recessive	8a, 8b	M, H, R, (C)	METHYLMALONIC ACIDURIA AND HOMOCYSTINURIA
MMADHC	Mendelian autosomal	recessive	5	M, H, R, E	METHYLMALONIC ACIDURIA AND HOMOCYSTINURIA
MOCS1	Mendelian autosomal	recessive	2	M, H, G, C, E	MOLYBDENUM COFACTOR DEFICIENCY
MOCS2	Mendelian autosomal	recessive	2	M, H, G, C, E	MOLYBDENUM COFACTOR DEFICIENCY
MPDU1	Mendelian autosomal	recessive	2	M, O, S	CONGENITAL DISORDER OF GLYCOSYLATION
MPLKIP	Mendelian autosomal	recessive	5	A, S, H, K	TRICHOTHIODYSTROPHY
MT-ATP6	Mitochondrial	N/A	8b	M, H, T, Q, G	NEUROPATHY, ATAXIA, AND RETINITIS PIGMENTOSA
MT-CO2	Mitochondrial	N/A	8b	M, H, G, C, Q	CYTOCHROME c OXIDASE

					DEFICIENCY
MT-CO3	Mitochondrial	N/A	8b	M, H, G,Q	MITOCHONDRIAL COMPLEX IV DEFICIENCY
MT-COI	Mitochondrial	N/A	8b	M, H, E	CYTOCHROME c OXIDASE I DEFICIENCY
MT-ND5	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q	Leigh syndrome
MT-ND5	Mitochondrial	N/A	8b	M, L, E, H, G, Q	MELAS
MTR	Mendelian autosomal	N/A	8b	M, H	METHYLCOBALAMIN DEFICIENCY
MTRR	Mendelian autosomal	recessive	8b	M, H	HOMOCYSTINURIA-MEGALOBLASTIC ANEMIA DUE TO DEFECT IN COBALAMIN METABOLISM
MT-TK	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q	Leigh syndrome
MT-TK	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q	MELAS syndrome
MT-TK	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q	MERFF syndrome
MT-TL1	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q, A	MELAS syndrome
MT-TL1	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q, A	Leigh syndrome
MT-TS1	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q	mitochondrial complex IV deficiency
MT-TV	Mitochondrial	N/A	8b	M, L, E, H, G, C, Q	Leigh syndrome
MUT	Mendelian autosomal	recessive	8a, 8b	M, H, W, A	METHYLMALONIC ACIDURIA
MVK	Mendelian autosomal	recessive	5	M, R, H, J, E, (C)	MEVALONIC ACIDURIA
MYCN	Mendelian autosomal	dominant	4	B, X, Ua, V	FEINGOLD SYNDROME
MYO5A	Mendelian autosomal	recessive	5	H, S	GRISCELLI SYNDROME
NAGA	Mendelian autosomal	recessive	5	M, H, G, S	SCHINDLER DISEASE
NAGA	Mendelian autosomal	recessive	5	M, H, G, S	KANZAKI DISEASE
NAGLU	Mendelian autosomal	recessive	8b	M, G, P, C	MUCOPOLYSACCHARIDOSIS
NBN	Mendelian autosomal	recessive	5	A, B, I, J, K	NIJMEGEN BREAKAGE SYNDROME
NDP	Mendelian X-linked	recessive	8a	T, P	NORRIE DISEASE
NDUFA1	Mendelian X-linked	not sure	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX I DEFICIENCY
NDUFA11	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX I DEFICIENCY
NDUFS1	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX I DEFICIENCY
NDUFS2	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX I DEFICIENCY
NDUFS3	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX I DEFICIENCY
NDUFS3	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	LEIGH SYNDROME
NDUFS4	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	LEIGH SYNDROME
NDUFS4	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX I DEFICIENCY
NDUFS7	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	LEIGH SYNDROME
NDUFS8	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	LEIGH SYNDROME
NDUFV1	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX I DEFICIENCY
NEU1	Mendelian autosomal	recessive	5	M, H, G, U	NEURAMINIDASE DEFICIENCY
NF1	Mendelian autosomal	dominant	8a	I, S, T	NEUROFIBROMATOSIS
NF1	Mendelian autosomal	dominant	8a	I, S, T	NEUROFIBROMATOSIS-NOONAN SYNDROME
NF1	Mendelian autosomal	dominant	5	I, S, T	WATSON SYNDROME
NHS	Mendelian X-linked	not sure	7	T, X	NANCE-HORAN SYNDROME
NIPBL	Mendelian autosomal	dominant	1	A, B, Ua, V, P, W	CORNELIA DE LANGE SYNDROME
NKX2-1	Mendelian autosomal	dominant	8a	K, H	CHOREOATHETOSIS, HYPOTHYROIDISM, AND NEONATAL RESPIRATORY DISTRESS
NLGN3	Mendelian X-linked	not sure	6	P, (E)	ASPERGER SYNDROME, X-LINKED, SUSCEPTIBILITY TO
NLGN3	Mendelian X-linked	not sure	6	P	AUTISM, SUSCEPTIBILITY TO
NLGN4X	Mendelian X-linked	not sure	6	P	AUTISM, SUSCEPTIBILITY TO
NLRP3	Mendelian autosomal	dominant	8a	J, H, S, G, (A)	CINCA SYNDROME
NPHP1	Mendelian autosomal	recessive	4	H, L, O, W	JOUBERT SYNDROME

NRXN1	Mendelian autosomal	recessive	2	O, E, P	PITT-HOPKINS-LIKE SYNDROME
NRXN1	Mendelian autosomal	dominant	6	P	ASD/autism
NSD1	Mendelian autosomal	dominant	4, 5	F, P, V, W, E, L, Ub	SOTOS SYNDROME
NSD1	Mendelian autosomal	dominant	4, 5	F, P, V, W, E, L, Ub	WEAVER SYNDROME
NTRK1	Mendelian autosomal	recessive	5	H, O, S, P	INSENSITIVITY TO PAIN, CONGENITAL, WITH ANHIDROSIS
OCRL	Mendelian X-linked	not sure	5	M, H, E, A, T, W, P	LOWE OCULOCEREBRORENAL SYNDROME
OCRL	Mendelian X-linked	not sure	5	M, A, W	DENT DISEASE
OFD1	Mendelian X-linked	dominant	4	L, Ua, Uc, W	OROFACIODIGITAL SYNDROME
OFD1	Mendelian X-linked	recessive	4	L, Ua, T, W, (H, O)	JOUBERT SYNDROME
OFD1	Mendelian X-linked	recessive	1	W, F, Ua, Uc, S, V, X, H	SIMPSON-GOLABI-BEHMEL SYNDROME
OPHN1	Mendelian X-linked	recessive	1	L, H, E	MENTAL RETARDATION, X-LINKED
PAFAH1B1	Mendelian autosomal	dominant	1	Ba, E, L, C, H	LISSENCEPHALY
PAK3	Mendelian X-linked	recessive	6	(B), P	MENTAL RETARDATION, X-LINKED
PANK2	Mendelian autosomal	recessive	8b	H, G, L, T, C, P	NEURODEGENERATION WITH BRAIN IRON ACCUMULATION
PANK2	Mendelian autosomal	recessive	8b	H, G, L, T, M	HYPOPREBETALIPOPROTEINEMIA, ACANTHOCYTOSIS, RETINITIS PIGMENTOSA, AND PALLIDAL DEGENERATION
PAX6	Mendelian autosomal	dominant	4	T, H, (L)	ANIRIDIA, CEREBELLAR ATAXIA, AND MENTAL RETARDATION
PAX6	Mendelian autosomal	dominant	7	T, (A)	COLOBOMA OF OPTIC NERVE
PAX6	Mendelian autosomal	dominant	7	T, (L)	ANIRIDIA
PC	Mendelian autosomal	recessive	5	M, C, E, H, G, L	PYRUVATE CARBOXYLASE DEFICIENCY
PCNT	Mendelian autosomal	recessive	8a	AA, B, U	MICROCEPHALIC OSTEODYSPLASTIC PRIMORDIAL DWARFISM
PDHA1	Mendelian X-linked	not sure	8b	M, H, G, E, C, L	PYRUVATE DEHYDROGENASE E1-ALPHA DEFICIENCY
PDHA1	Mendelian X-linked	not sure	8b	M, H, G, E, C, L	LEIGH SYNDROME
PDSS1	Mendelian autosomal	recessive	8b	M, H, T, O, V	COENZYME Q10 DEFICIENCY, PRIMARY
PDSS2	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, W	COENZYME Q10 DEFICIENCY, PRIMARY
PEPD	Mendelian autosomal	recessive	5	M, J, S	PROLIDASE DEFICIENCY
PEX1	Mendelian autosomal	recessive	2	M, E, H, T, C, L	PEROXISOME BIOGENESIS DISORDER (ZELLWEGER)
PEX1	Mendelian autosomal	recessive	5	M, H, T, G	PEROXISOME BIOGENESIS DISORDER
PEX10	Mendelian autosomal	recessive	2	M, E, H, T, G, C, L	PEROXISOME BIOGENESIS DISORDER (ZELLWEGER)
PEX10	Mendelian autosomal	recessive	5	M, E, H, T, G, C	PEROXISOME BIOGENESIS DISORDER
PEX12	Mendelian autosomal	recessive	2	M, E, H, T, G, C, L	PEROXISOME BIOGENESIS DISORDER (ZELLWEGER)
PEX12	Mendelian autosomal	recessive	2	M, E, H, T, G, C, L	PEROXISOME BIOGENESIS DISORDER
PEX13	Mendelian autosomal	recessive	2	M, E, H, T, G, C, L	PEROXISOME BIOGENESIS DISORDER (ZELLWEGER)
PEX13	Mendelian autosomal	recessive	2	M, E, H, T, G, C	PEROXISOME BIOGENESIS DISORDER
PEX26	Mendelian autosomal	recessive	8b	M, E, H, L, C, (G)	PEROXISOME BIOGENESIS DISORDER (ZELLWEGER)
PEX26	Mendelian autosomal	recessive	8b	M, E, H, L, C, (G)	PEROXISOME BIOGENESIS DISORDER
PEX5	Mendelian autosomal	recessive	2	M, E, H, T, G, C, L	PEROXISOME BIOGENESIS DISORDER (ZELLWEGER)
PEX5	Mendelian autosomal	recessive	2	M, E, H, T, G, C	PEROXISOME BIOGENESIS DISORDER
PEX6	Mendelian autosomal	recessive	2	M, E, H, T, G, C	PEROXISOME BIOGENESIS DISORDER (ZELLWEGER)
PEX6	Mendelian autosomal	recessive	2	M, T, H, C, G	PEROXISOME BIOGENESIS DISORDER
PEX7	Mendelian autosomal	recessive	1	AAa, M, E, T, C, Ub	RHIZOMELIC CHONDRODYSPLASIA PUNCTATA
PEX7	Mendelian autosomal	recessive	5	M, H, T, G	PEROXISOME BIOGENESIS DISORDER

PGK1	Mendelian X-linked	recessive	8a	H, Q, R, M	PHOSPHOGLYCERATE KINASE 1 DEFICIENCY
PHF6	Mendelian X-linked	recessive	5	O, A, K	BORJESON-FORSSMAN-LEHMANN SYNDROME
PHF8	Mendelian X-linked	recessive	4	Uc	SIDERIUS X-LINKED MENTAL RETARDATION SYNDROME
PHGDH	Mendelian autosomal	recessive	5	M, Bc, E	PHOSPHOGLYCERATE DEHYDROGENASE DEFICIENCY
PLP1	Mendelian X-linked	recessive	2	H, L, E, (C)	PELIZAEUS-MERZBACHER DISEASE
PMM2	Mendelian autosomal	recessive	5	M, H, L, T, A, C, R	CONGENITAL DISORDER OF GLYCOSYLATION
PNP	Mendelian autosomal	recessive	8a	M, J, H	PURINE NUCLEOSIDE PHOSPHORYLASE DEFICIENCY
POMGNT1	Mendelian autosomal	recessive	1	Q, L, T, C, E, P, G	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH BRAIN AND EYE ANOMALIES)
POMGNT1	Mendelian autosomal	recessive	4	Q, L, T	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH MENTAL RETARDATION)
POMT1	Mendelian autosomal	recessive	1	Q, L, T, B, C	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH BRAIN AND EYE ANOMALIES)
POMT1	Mendelian autosomal	recessive	4	Q, L, T, C	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH MENTAL RETARDATION)
POMT2	Mendelian autosomal	recessive	1	Q, L, C, T, B	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH BRAIN AND EYE ANOMALIES)
POMT2	Mendelian autosomal	recessive	1	Q, L, C, B, T	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (CONGENITAL WITH MENTAL RETARDATION)
PORCN	Mendelian X-linked	dominant	7	S, T, Ua, Uc, X	FOCAL DERMAL HYPOPLASIA
PPOX	Mendelian autosomal	not sure	8a	M, H, S, E	PORPHYRIA VARIEGATA
PQBP1	Mendelian X-linked	recessive	1	A, B, W, X	RENPENNING SYNDROME
PRPS1	Mendelian X-linked	recessive	5	H, J, C, M	ARTS SYNDROME
PRSS12	Mendelian autosomal	recessive	6		MENTAL RETARDATION, AUTOSOMAL RECESSIVE
PTCH1	Mendelian autosomal	dominant	4	L, Uc	HOLOPROSENCEPHALY
PTEN	Mendelian autosomal	dominant	8a	F, P, (I)	MACROCEPHALY/AUTISM SYNDROME
PTPN11	Mendelian autosomal	dominant	7	A, V, R, T, S, Q	NOONAN SYNDROME
PTPN11	Mendelian autosomal	dominant	7	A, V, R, T, S, Q	LEOPARD SYNDROME
PUS1	Mendelian autosomal	recessive	8a	M, Q, R	MYOPATHY, LACTIC ACIDOSIS, AND SIDEROBLASTIC ANEMIA
PVRL1	Mendelian autosomal	recessive	7	Uac, S	CLEFT LIP/PALATE-ECTODERMAL DYSPLASIA SYNDROME
PYCR1	Mendelian autosomal	recessive	5	A, S, (L), (B), U	CUTIS LAXA, AUTOSOMAL RECESSIVE
PYCR1	Mendelian autosomal	recessive	5	A, S, T, U	CUTIS LAXA, AUTOSOMAL RECESSIVE
PIGV	Mendelian autosomal	recessive	1	M, Ua, (E)	HYPERPHOSPHATASIA WITH MENTAL RETARDATION SYNDROME
RAB27A	Mendelian autosomal	recessive	8a	H, R, J, S, C	GRISCELLI SYNDROME
RAB3GAP1	Mendelian autosomal	recessive	1	B, T, L, H	WARBURG MICRO SYNDROME
RAB3GAP2	Mendelian autosomal	recessive	4	B, T, W, H	MARTSOLF SYNDROME
RAB3GAP2	Mendelian autosomal	recessive	1	T, H, B, L, X	WARBURG MICRO SYNDROME
RAB39B	Mendelian X-linked	not sure	6	(E), (P), (F)	MENTAL RETARDATION, X-LINKED
RAF1	Mendelian autosomal	dominant	7	A, V, R, T, Q	NOONAN SYNDROME
RAF1	Mendelian autosomal	dominant	7	A, V, R, T, Q	LEOPARD SYNDROME
RAI1	Mendelian autosomal	dominant	5	P, H, J	SMITH-MAGENIS SYNDROME
RBM28	Mendelian autosomal	recessive	2	K, S, H, G	ALOPECIA, NEUROLOGIC DEFECTS, AND ENDOCRINOPATHY SYNDROME
RELN	Mendelian autosomal	recessive	1	L, E, H	LISSENCEPHALY
RFT1	Mendelian autosomal	recessive	2	H, E, B, M	CONGENITAL DISORDER OF GLYCOSYLATION
RMRP	Mendelian autosomal	recessive	8a	AA, B, U	ANAUXETIC DYSPLASIA
RNASEH2A	Mendelian autosomal	recessive	2	H, E, L, S, G, C, B	AICARDI-GOUTIERES SYNDROME

RNASEH2B	Mendelian autosomal	recessive	2	H, E, L, S, G, B	AICARDI-GOUTIERES SYNDROME
RNASEH2C	Mendelian autosomal	recessive	2	H, E, L, S, G, C, B	AICARDI-GOUTIERES SYNDROME
RPGRIP1L	Mendelian autosomal	recessive	4	H, L, O, W, T, Ua	JOUBERT SYNDROME
RPGRIP1L	Mendelian autosomal	recessive	4	H, L, O, W, T	COACH SYNDROME
RPL10	Mendelian X-linked	not sure	6	P	AUTISM, SUSCEPTIBILITY TO
RPS6KA3	Mendelian X-linked	not sure	2	A, B, Ub, (H), (C)	COFFIN-LOWRY SYNDROME
RPS6KA3	Mendelian X-linked	not sure	6		MENTAL RETARDATION, X-LINKED
SALL1	Mendelian autosomal	dominant	7	X, Ua, V, W	TOWNES-BROCKS SYNDROME
SATB2	Mendelian autosomal	dominant	1	E, Uc	CLEFT PALATE, ISOLATED
SC5DL	Mendelian autosomal	recessive	4	M, B, Ua, L	LATHOSTEROLOSIS
SCO2	Mendelian autosomal	recessive	8b	M, L, H, G, C, Q	CARDIOENCEPHALOMYOPATHY, FATAL INFANTILE, DUE TO CYTOCHROME c OXIDASE DEFICIENCY
SDHA	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	LEIGH SYNDROME
SDHA	Mendelian autosomal	recessive	8b	M, L, E, H, G, C, Q	MITOCHONDRIAL COMPLEX II DEFICIENCY
SETBP1	Mendelian autosomal	dominant	1	U, V, W, I, X, C	SCHINZEL-GIEDION MIDFACE RETRACTION SYNDROME
SETBP1	Mendelian autosomal	dominant	5	(P)	unspecific moderate ID
SHANK2	Mendelian autosomal	dominant	6	P	AUTISM, SUSCEPTIBILITY TO
SHANK3	Mendelian autosomal	dominant	6	P	CHROMOSOME 22q13.3 DELETION SYNDROME
SHH	Mendelian autosomal	dominant	4	L, Uc, C	HOLOPROSENCEPHALY
SHH	Mendelian autosomal	dominant	4	L	SCHIZENCEPHALY
SHOC2	Mendelian autosomal	dominant	4	A, S, V, (Q)	NOONAN SYNDROME-LIKE DISORDER WITH LOOSE ANAGEN HAIR
SHROOM4	Mendelian X-linked	recessive	2	A, U	STOCCO DOS SANTOS X-LINKED MENTAL RETARDATION SYNDROME
SIL1	Mendelian autosomal	recessive	5	H, L, Q, T, A, Ub	MARINESCO-SJOGREN SYNDROME
SIX3	Mendelian autosomal	dominant	4	L, C	HOLOPROSENCEPHALY
SIX3	Mendelian autosomal	dominant	4	L, C	SCHIZENCEPHALY
SLC12A6	Mendelian autosomal	recessive	4	L, G, H, Ub, P	AGENESIS OF THE CORPUS CALLOSUM WITH PERIPHERAL NEUROPATHY
SLC16A2	Mendelian X-linked	not sure	2	K, H, Q, P, E, L, (A), (B)	ALLAN-HERNDON-DUDLEY SYNDROME
SLC17A5	Mendelian autosomal	recessive	8b	M, H, G, E, C, L	INFANTILE SIALIC ACID STORAGE DISORDER
SLC17A5	Mendelian autosomal	recessive	8b	M, H, G, E, L	SIALURIA
SLC25A15	Mendelian autosomal	recessive	5	M, H, E	HYPERORNITHINEMIA-HYPERAMMONEMIA-HOMOCITRULLINURIA SYNDROME
SLC25A22	Mendelian autosomal	recessive	2	E, B, H, L	EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE
SLC2A1	Mendelian autosomal	dominant	5	M, E, H	GLUT1 DEFICIENCY SYNDROME
SLC2A1	Mendelian autosomal	dominant	5	M, E, H, B	GLUT1 DEFICIENCY SYNDROME
SLC2A1	Mendelian autosomal	dominant	5	M, E, H	DYSTONIA
SLC35C1	Mendelian autosomal	recessive	2	M, A, B, H, J	CONGENITAL DISORDER OF GLYCOSYLATION
SLC4A4	Mendelian autosomal	recessive	5	A, T, W	RENAL TUBULAR ACIDOSIS, PROXIMAL, WITH OCULAR ABNORMALITIES AND MENTAL RETARDATION
SLC6A8	Mendelian X-linked	not sure	5	M, H, E, P	CREATINE DEFICIENCY SYNDROME
SMC1A	Mendelian X-linked	not sure	1	A, B, V, E, X	CORNELIA DE LANGE SYNDROME
SMC3	Mendelian autosomal	dominant	5		ORNELIA DE LANGE SYNDROME
SMPD1	Mendelian autosomal	recessive	8b	M, C, G, H, J, T	NIEMANN-PICK DISEASE
SMPD1	Mendelian autosomal	recessive	5	M, A, C, H, J, T	NIEMANN-PICK DISEASE
SMS	Mendelian X-linked	recessive	5	Ub, H, E	MENTAL RETARDATION, X-LINKED, SYNDROMIC
SNAP29	Mendelian autosomal	recessive	2	B, H, L, S	CEREBRAL DYSGENESIS, NEUROPATHY, ICHTHYOSIS, AND PALMOPLANTAR KERATODERMA SYNDROME
SOS1	Mendelian autosomal	dominant	7	V, T	NOONAN SYNDROME
SOX10	Mendelian autosomal	dominant	5	S, X, (L), H	WAARDENBURG SYNDROME
SOX10	Mendelian autosomal	dominant	5	S, X, L, H	PERIPHERAL DEMYELINATING NEUROPATHY, CENTRAL DYSMYELINATION, WAARDENBURG SYNDROME, AND HIRSCHSPRUNG

					DISEASE
SOX2	Mendelian autosomal	dominant	4	L, T, A, W, X	MICROPHTHALMIA, SYNDROMIC
SOX3	Mendelian X-linked	not sure	5	A, K	MENTAL RETARDATION, X-LINKED, WITH PANHYPOPITUITARISM
SPRED1	Mendelian autosomal	dominant	8a	S, (P)	LEGIUS SYNDROME
SRPX2	Mendelian X-linked	not sure	5	E, H	ROLANDIC EPILEPSY, MENTAL RETARDATION, AND SPEECH DYSPRAXIA, X-LINKED
ST3GAL3	Mendelian autosomal	recessive	6		MENTAL RETARDATION, AUTOSOMAL RECESSIVE
ST3GAL3	Mendelian autosomal	recessive	2	E, H	EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE
STIL	Mendelian autosomal	recessive	5	B	MICROCEPHALY, PRIMARY, AUTOSOMAL RECESSIVE
STRA6	Mendelian autosomal	recessive	1	C, T, V, X	MICROPHTHALMIA, SYNDROMIC
SUOX	Mendelian autosomal	recessive	5	T, H, M, (G), (C)	SULFOCYSTEINURIA
SURF1	Mendelian autosomal	recessive	8b	M, L, E, H, G, Q	LEIGH SYNDROME
SYN1	Mendelian X-linked	recessive	8a	E, P	EPILEPSY, X-LINKED, WITH VARIABLE LEARNING DISABILITIES AND BEHAVIOR DISORDERS
SYNGAP1	Mendelian autosomal	dominant	3	E	MENTAL RETARDATION, AUTOSOMAL DOMINANT
SYP	Mendelian X-linked	not sure	6	(E)	MENTAL RETARDATION, X-LINKED
TAT	Mendelian autosomal	recessive	8a	M, S, T	TYROSINEMIA
TBCE	Mendelian autosomal	recessive	5	A, K, U	HYPOPARATHYROIDISM-RETARDATION-DYSMORPHISM SYNDROME
TBCE	Mendelian autosomal	recessive	5	A, K, U, J	KENNY-CAFFEY SYNDROME
TCF4	Mendelian autosomal	dominant	2	E, O, T, B	PITT-HOPKINS SYNDROME
TGFBR1	Mendelian autosomal	dominant	7	Ub, Uc, (S)	LOEYS-DIETZ SYNDROME
TGFBR2	Mendelian autosomal	dominant	7	Ub, Uc, (S)	LOEYS-DIETZ SYNDROME
TGIF1	Mendelian autosomal	dominant	4	L, Uc, C, X	HOLOPROSENCEPHALY
THRB	Mendelian autosomal	dominant	8a	K, (A)	THYROID HORMONE RESISTANCE, GENERALIZED, AUTOSOMAL DOMINANT
TIMM8A	Mendelian X-linked	recessive	8b	H, G, T, P	MOHR-TRANEBJAERG SYNDROME
TMEM67	Mendelian autosomal	recessive	4	H, L, W, T	JOUBERT SYNDROME
TMEM67	Mendelian autosomal	recessive	4	H, L, W, T	COACH SYNDROME
TPI1	Mendelian autosomal	recessive	8a	M, R, H, C	"TPI deficiency"
TREX1	Mendelian autosomal	recessive	2	H, E, L, S, G, C, B	AICARDI-GOUTIERES SYNDROME
TRIM32	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
TSC1	Mendelian autosomal	dominant	5	E, I, L, S, W	TUBEROUS SCLEROSIS
TSC1	Mendelian autosomal	dominant	5	E, L	FOCAL CORTICAL DYSPLASIA OF TAYLOR
TSC2	Mendelian autosomal	dominant	5	E, I, L, S, W	TUBEROUS SCLEROSIS
TSPAN7	Mendelian X-linked	recessive	6		MENTAL RETARDATION, X-LINKED
TTC8	Mendelian autosomal	recessive	4	N, T, Ua, W, K, P	Bardet-Biedl syndrome
TUBA1A	Mendelian autosomal	dominant	1	L, B, H, E	LISSENCEPHALY
TUBB2B	Mendelian autosomal	dominant	1	L, E, H	POLYMICROGYRIA
TUSC3	Mendelian autosomal	recessive	3		MENTAL RETARDATION, AUTOSOMAL RECESSIVE
UBE2A	Mendelian X-linked	recessive	2	F, W, E	MENTAL RETARDATION, X-LINKED, SYNDROMIC
UBE3A	Mendelian autosomal	dominant	2	E, B, H, P	ANGELMAN SYNDROME
UBR1	Mendelian autosomal	recessive	4	X, K, S, (W), (A), (B)	JOHANSON-BLIZZARD SYNDROME
UPB1	Mendelian autosomal	recessive	5	M, H	BETA-UREIDOPROPIONASE DEFICIENCY
UPF3B	Mendelian X-linked	recessive	5	P, Q	MENTAL RETARDATION, X-LINKED, SYNDROMIC
VLDLR	Mendelian autosomal	recessive	2	E, H, L, (T), (A)	CEREBELLAR ATAXIA, MENTAL RETARDATION, AND DYSEQUILIBRIUM SYNDROME
VPS13B	Mendelian autosomal	recessive	2	B, N, T, I, R, A, K	COHEN SYNDROME
XPA	Mendelian autosomal	recessive	8a, 8b	B, H, G, S, T, I, (C)	XERODERMA PIGMENTOSUM
ZC3H14	Mendelian autosomal	recessive	6		non syndromic autosomal recessive ID
ZDHHC9	Mendelian X-linked	recessive	2	(U)	MENTAL RETARDATION, X-LINKED, SYNDROMIC
ZEB2	Mendelian autosomal	dominant	1	B, E, V, W, X, A, L	MOWAT-WILSON SYNDROME
ZFYVE26	Mendelian autosomal	recessive	5	H, G, (L)	SPASTIC PARAPLEGIA, AUTOSOMAL

					RECESSIVE
ZIC2	Mendelian autosomal	dominant	4	L, B	HOLOPROSENCEPHALY
ZNF41	Mendelian X-linked	not sure	6	P	MENTAL RETARDATION, X-LINKED
ZNF674	Mendelian X-linked	recessive	3		MENTAL RETARDATION, X-LINKED
ZNF711	Mendelian X-linked	not sure	6		MENTAL RETARDATION, X-LINKED
ZNF81	Mendelian X-linked	not sure	6	(A)	MENTAL RETARDATION, X-LINKED

List of 388 ID genes (status 2010) that was used for the enrichment analyses. This short version of the original catalog contains data on inheritance mode, associated ID-disorders, main class and additional phenotypes.

Chapter 7

General Discussion

In this thesis haploinsufficiency of *TCF4* as the underlying cause for Pitt-Hopkins syndrome was identified (**chapter 2**). Subsequently, recessive defects in *NRXN1* and *CNTNAP2* were detected in patients with a similar clinical phenotype (**chapter 3**). These findings and the associated molecular and functional studies helped to delineate the clinical phenotype on the one hand (**chapters 2-5**) and to provide inside into the (interconnected) molecular mechanisms on the other hand (**chapters 2 and 4**). Furthermore, with the comprehensive catalogue of all currently known ID genes and their associated phenotypes (**chapter 6**), a resource for future systematic studies in this field was generated, and first global insights were obtained.

7.1. From phenotype to genotype and back

For decades genetic diagnosis and assignment to particular syndromes was based on evaluation and description of the clinical phenotype. Only when the underlying genetic defects were identified, these clinical groups could be confirmed or not. The increasing molecular possibilities resulted in a fusion of phenotypic analysis with molecular biology and a new lumping and splitting strategy, taking full advantage of both disciplines.[330] This study is an example how close phenotypic and molecular findings can interact and complement each other.

Pitt-Hopkins syndrome (PTHS) was initially described in 1978 in two unrelated patients with a supposedly distinct phenotype of severe ID, hyperbreathing and recognizable facial gestalt with a beaked nose.[227] However, in the following nearly 30 years only a few single case reports described patients with a possibly identical disorder.[228-231] Only after the identification of *TCF4* haploinsufficiency by us and others as the underlying cause of Pitt-Hopkins syndrome (**chapter 2**),[247,250] the number of clinically identified and subsequently molecularly confirmed patients increased rapidly (**chapter 3).**[260,261,331-333] Now the diagnosis could be confirmed or disproved in some of the previously described patients. This helped to delineate the clinical hallmarks of a distinct and recognizable clinical phenotype that apparently had been under-recognized for decades (**chapter 2**). The improved phenotypic delineation and the opportunity to perform confirmative molecular testing extensively raised the awareness for this syndrome amongst clinical geneticists and established it as a differential diagnosis in the field of severe ID together with Angelman- and Rett syndromes (**chapter 3**). To facilitate diagnosis even for less experienced clinicians, several groups recently attempted to establish clinical scores for Pitt-Hopkins syndrome.[331,333]

For a long time, gene identification in ID was mainly limited to familial cases, allowing linkage analysis or homozygosity mapping.[39,55,57] Identifying genes for sporadic ID disorders mostly relied on the random detection of chromosomal aberrations.[68,150] With the

development of molecular karyotyping by arrayCGH or SNP array a new tool to identify more and smaller of such chromosomal aberrations became available. By detecting small deletions in one patient and by subsequently identifying mutations in one of the deleted genes in other patients with the same disorder but without chromosomal anomalies, the underlying cause for several disorders could be identified.[36,73] Also in this study, molecular karyotyping confirmed its usefulness in disease gene identification by contributing to the identification of three underlying genetic causes for ID. D*e novo* defects in *TCF4* were identified in Pitt-Hopkins syndrome, and recessive defects in *NRXN1* or *CNTNAP2* were identified in patients with clinically overlapping disorders and no history of parental consanguinity (**chapter 2 and 4**). However, it also demonstrated, how valuable accurate clinical characterization is, as only identification of a *de novo* mutation in a second patient proved the deletion of *TCF4* being the specific underlying cause in the first patient **(chapter 2)**. Both patients had previously been described with a clinical diagnosis of Pitt-Hopkins syndrome.[229] In contrast, screening the gene in a larger group of patients with idiopathic ID but only mild phenotypic overlap did not lead to the identification of a second confirmatory patient.[250]

The approach to systematically perform molecular karyotyping in a larger group of clinically homogenous patients - all referred with suspected Pitt-Hopkins syndrome but without *TCF4* defect - lead to the identification of homozygous or compound heterozygous defects in *CNTNAP2* in a pair of siblings and a sporadic patient and in *NRXN1* in a sporadic patient. Their phenotype has been described as Pitt-Hopkins-like ID **(chapter 4)**. Interestingly, a homozygous stop mutation in *CNTNAP2* had previously been reported in patients with a cortical dysplasia focal epilepsy syndrome (CDFE), characterized by early onset of epilepsy and severe ID, following an initially normal development and subsequent regression.[184] The patients with recessive defects in *CNTNAP2* in our study also had early onset of epilepsy and normal or only mildly delayed early motor development. However, speech development was lacking from the beginning (**chapter 4**). While our patients had initially been diagnosed with Pitt-Hopkins syndrome due to severe ID with accompanying epilepsy, the patients published by Strauss et al.[184] had been diagnosed with an epilepsy syndrome. Common genotypes can therefore bring together phenotypes that were previously considered as distinct disorders, probably rather due to clinical bias than due to a reliable genotype-phenotype correlation.

Of note, all patients with recessive defects in *CNTNAP2* showed severe speech impairment. Interestingly, a molecular link between CNTNAP2 and FOXP2, a member of the forkhead transcription factor family had previously been shown.[281] Mutations in *FOXP2* are known to cause speech and language disorders.[334,335]

Identification of further patients with Pitt-Hopkins syndrome did not only result in a better delineation of the phenotype but also of the genotype **(chapter 2 and 3)**. The spectrum of aberrations in *TCF4* comprises deletions, truncating mutations, and missense mutations which are located in exons coding for the bHLH domain of TCF4. The fact that all of these defects resulted in a similar and recognizable phenotype indicated haploinsufficiency as the underlying mechanism. With an *in vitro* transcriptional reporter assay after co-transfection of wild type or mutant TCF4 with an ASCL1 expression construct it could be shown that the missense mutations in the DNA- and protein-binding bHLH domain of TCF4 resulted in loss of the transactivating capacities, thus validating *TCF4* haploinsufficiency as the underlying mechanism **(chapter 2)**.

ASCL1 is a tissue specific HLH transcription factor, knowing to interact with TCF4[236] and being part of the RET-signaling pathway. This additionally includes the PHOX2, RET and TH genes that are responsible for the development of transient or permanent noradrenergic derivatives.[237] Mutations in ASCL1 and PHOX2 can cause the congenital central hypoventilation syndrome (CCHS)[237,336] and mutations in RET are implicated in Hirschsprung disease.[337,338] Therefore, interaction of TCF4 with ASCL1 and the reduced transactivation capacities upon TCF4 mutations might explain some of the PTHS specific phenotypic aspects such as the breathing anomalies and severe constipation. The findings from the cell-based luciferase assay were confirmed later by a similar experiment in a neuronal cell line.[260]

Two recent studies investigated the consequences of different TCF4 missense mutations in various in vitro assays, finding evidence for a range of effects from hypomorphic to dominant-negative, thus contributing to the understanding of clinical variability in PTHS.[316,339] However, these effects are based on single assays, thus possibly explaining some of the clinical variability but not contradicting TCF4 haploinsufficiency as the general underlying mechanism.

7.2. From human to fly and back

During the screening of Pitt-Hopkins syndrome-like patients without *TCF4* mutation, recessive defects in *CNTNAP2* in a pair of siblings and an isolated patient as well as recessive defects in *NRXN1* in an isolated patient were detected (**chapter 4**). This was quite exciting as NRXN1 and CNTNAP2 belong to the same superfamily of neurexins, though no common molecular functions were known so far (see below). A potential common molecular basis and maybe even a link to TCF4 could therefore provide an explanation for the similar phenotype in the patients. However, neither *NRXN1* nor *CNTNAP2* is expressed in appreciable levels in blood, and other tissues are difficult to obtain from living patients.

Therefore *Drosophila melanogaster* was utilized as a model organism for further functional studies as all three genes do have orthologs in the fly: daughterless (TCF4), Nrx-I (NRXN1), and Nrx-IV (CNTNP2).

Our first straightforward hypothesis was that TCF4 as a transcription factor could regulate the expression levels of *NRXN1* and/or *CNTNAP2*. Ubiquitous knockdown of the TCF4 ortholog daughterless in the fly did not reveal conclusive data on altered expression levels of Nrx-I or Nrx-IV, the NRXN1 and CNTNAP2 orthologs, respectively, in the used system (**chapter 4**). However, recently another group reported a possible regulation of NRXN1 and CNTNAP2 through TCF4 in terms of a possible transactivation effect in cell system based promoter studies.[316]

NRXN1/Nrx-I has been known to be one of the key synapse organizing molecules by forming bridges over the synaptic cleft with its postsynaptic binding partners, the neuroligins.[287] Vertebrate CNTNAP2, also termed CASPR2 has been mainly known for its role in regulating neuron-glia contact and for colocalizing with K+- channels in the juxtaparanodal areas of Ranvier nodes in myelinated axons of both the central and peripheral nervous system.[294,295] When this study was initiated, fly Nrx-IV was reported to be almost exclusively expressed in glia-cells and to regulate glia-glia contact.[293,296]

The executed *Drosophila* experiments in this study brought several new insights (**chapter 4**) regarding NRXN1/Nrx-I and CNTNAP2/Nrx-IV and their possible interaction.

1) Nrx-IV was shown to play a role in neurons as neuronal knockdown lead to embryonic lethality. Previously, Nrx-IV was only considered to be expressed in glia cells in the fly.[296] In parallel, two groups reported on neuronal Nrx-IV isoforms,[264,268] thus supporting our conclusion that Nrx-IV plays a crucial role in neurons.

2) Nrx-IV is present at synapses. Previously, only a report on detection of Caspr2 in fractionated rat synaptic plasma had pointed to a possible synaptic presence.[266]

3) Nrx-I and Nrx-IV converge on synaptic active zone protein bruchpilot (brp) as a common target. Adding to the observation of decreased brp levels in Nrx-I mutants,[270] our work showed that both Nrx-I or Nrx-IV levels determine brp levels bidirectionally and that either Nrx-I or Nrx-IV overexpression induces identical changes in synaptic morphology.

Bruchpilot is a peripheral membrane protein being located at the presynaptic active zones, the location of neurotransmitter release into the synaptic cleft. Through its PDZ domain binding site it is involved in a large complex with other active zone proteins.[298] Of note, also Nrx-I and Nrx-IV contain C-terminal PDZ-domain binding sites, therefore raising the possibilty that all three proteins are assembled into one synaptic complex or macromolecular network. Brp shows high sequence and functional homology to the vertebrate family of ELKS/CAST proteins,[298] corresponding to the human synaptic proteins ERC1 and ERC2.

Summarizing, these findings in *Drosophila* helped to gain more insight into function of Nrx-I and Nrx-IV in the nervous system and particularly implicated Nrx-IV in synapse biology. It also helped to establish a functional link between Nrx-I and Nrx-IV as a possible explanation for the similar human phenotypes. Furthermore, with brp orthologs ERC1 and ERC2 it produced promising candidate genes to go back to humans and screen these genes for mutations in patients with a PTHS-like phenotype.

Therefore, in this study *Drosophila* proved to be a valuable model organism for functional testing in ID. However, it is important to know what to expect from an animal model like *Drosophila*. It is no equivalent to human patients, but the same is applicable for all other animal models or *in vitro* assays. It cannot provide final evidence for the pathogenicity of particular mutations, but it can provide functional support for the role of particular genes in cognitive and behavioral function and dysfunction. As successfully demonstrated in this thesis, it can help to establish molecular networks. This is in particularly interesting as the theme of overlapping phenotypes being caused by genes that are linked to each other in molecular networks[178] gets more and more evident and attractive.

Drosophila is not the answer to all our questions but it is an extremely suitable and valuable tool to efficiently add pieces to the puzzle that has been keeping and will keep scientists with an interest in ID - geneticists, and neurobiologists - busy for a long time.

7.3. From mono- to oligogenic and back?

A common synaptic link between NRXN1 and CNTNAP2 and possibly TCF4 might result in the implication of all three encoding genes in disorders across diagnostic boundaries, severe intellectual disability at one end of the spectrum[184,247,248,285,308,340,341] and neuropsychiatric disorders with no or only mild cognitive impairment at the other end.[266,272-278,280-283,299-307,342,343]

Biallelic defects in *NRXN1* or *CNTNAP2* were observed in patients with severe cognitive disorders (**chapter 4**).[184,285,340] Heterozygous copy number variants (CNVs) and single nucleotide changes in either gene at that time had been found in patients with neuropsychiatric disorders and no or only mild cognitive impairment.[266,272-283,286,299-307] The apparent influence of gene-dosage on the phenotypic severity has been also perfectly mirrored in a published family. The index patient harboring a compound heterozygous deletion and splice site mutation in *NRXN1* has severe ID with epilepsy and little social interaction, while some of his family members carrying only one of the *NRXN1* aberrations have schizophrenia, psychotic disorders or sub-diagnostic autistic traits.[344]

The observation of heterozygous variants as risk factors for low penetrant neuropsychiatric disorders and of biallelic defects in the same gene as causative for fully penetrant severe ID has not been reported for other genes so far. Only recently, a

homozygous mutation in the *ANK3* gene was reported in a consanguineous family with moderate ID and severe behavioral anomalies.[345] *De novo*, heterozygous mutations in *ANK3* had previously been implicated in ASD, and polymorphisms had been implicated as susceptibility factors for psychiatric disorders. However, in this study a correlation between severity and degree of inactivation due to the mutation was discussed rather than dosage sensitivity.[345]

We continued to screen *NRXN1* and *CNTNAP2* in a larger cohort of patients with Pitt-Hopkins like syndrome with severe ID and accepted samples from external patients with severe ID and a heterozygous CNV in *NRXN1* or *CNTNAP2* for mutational screening of the second allele. In this group we expected to identify more patients with biallelic defects in either gene and therefore to further delineate the associated clinical phenotype (**chapter 5**). However, surprisingly, in none of seven patients with severe ID and a heterozygous *CNTNAP2* defect and in none of four patients with severe ID and a heterozygous *NRXN1* defect we could identify a mutation on the respective second allele. In most of the patients, the heterozygous defect was inherited from a healthy parent, therefore a second defect on the other allele would have been expected in order to explain the severe phenotype. Prompted by the molecular link between NRXN1 and CNTNAP2/CASPR2 fly orthologs (**chapter 4**) and considering a possible digenic pathomechanism, screening of the respective other gene was performed in these patients but without resulting in the detection of mutations.

Of course the second mutation in either gene might have remained undetected due to localization in a non-coding regulatory element or in a non-tested alternative isoform, or the finding of the *NRXN1* or *CNTNAP2* defect might have been just coincidence without pathogenic relevance to the severe phenotype of these patients. This might be true for the patient with a splice site variant in *CNTNAP2* and additionally a clearly pathogenic mutation in *MEF2C*.[346] However, the large number of such severely affected patients, further raised by recent reports at least for *NRXN1*,[347-349] might indicate the presence of a second and maybe third contributing factor somewhere else in the genome in terms of a digenic or oligogenic cause.

While mild forms of ID are assumed to represent the lower end of normal IQ distribution and to result from the interaction of various genetic and other factors, for severe ID mainly a single genetic cause is assumed.[2] This is confirmed by recent exome sequencing studies identifying *de novo* mutations in a large proportion of patients with unspecific severe ID.[34,35,42] For a small subset of developmental disorders that are frequently associated with mild cognitive impairment, such as autism or dyslexia, polygenic effects are discussed.[314,350] However, in contrast to the multifactorial “common disease, common variant” model, suggesting many genetic and non-genetic factors (epigenetic and environmental) with small

effect in individual patients,[351] the picture now rather points to single events of large effect. Recent studies on common neurodevelopmental disorders like autism spectrum disorders[47,116,350,352,353] or schizophrenia[354,355] revealed a large number of *de novo* mutations. One study on autism spectrum disorders estimated a 5- to 20-fold increased risk by spontaneous coding mutations in any of a large number of genes.[350] The concurrent finding of *de novo* defects in *SHANK2* with other inherited CNVs known to be risk factors for neuropsychiatric disorders would be in line with a multiple hit model for autism spectrum disorders.[46] This is also supported by the finding of a *de novo* frameshift mutation in *FOXP1* in combination with an inherited mutation in *CNTNAP2* in a patient with severe language delay, moderate ID and regression.[352] Before, *FOXP1* defects were found in patients with mild to moderate ID and language defects.[356,357] Therefore, a combination of *FOXP1* haploinsufficiency with increased *CNTNAP2* expression might contribute to the more severe phenotype in this patient.[352] Interestingly, for FOXP2, another member of the forkhead transcription factor family and implicated in language disorders, a molecular link to CNTNAP2 had previously been shown.[281]

While monogenic or single chromosomal defects might still be responsible for the majority of severe ID, there are already a few examples for disorders with digenic inheritance or a two hit model. For some cases of Bardet-Biedl syndrome, a di- or even trigenic inheritance has been reported.[48] Regarding a recurrent 16p12.1 microdeletion, second hits were shown to be necessary to evoke the full, severe phenotype.[45]

Only recently, two patients with heterozygous *NRXN1* deletions and additionally a *NRXN3* or a 16p11.2 deletion, respectively, were reported.[349] As defects in *NRXN3* as well the common *16p11.2* deletion have already been implicated in autism spectrum disorders,[342,358,359] these findings support the presence and absence of additional genetic lesions as contributory to the variable expressivity and incomplete penetrance in carriers of a *NRXN1* heterozygous exonic deletion.[349] By comparing the *de novo* rate of mutations and the burden of “second hits” in individuals with *NRXN1* deletions and carriers of known recurrent “susceptibility CNVs” the authors place *NRXN1* exonic deletions somewhere in the spectrum between 15q13.3 and 16p12.1 deletions.[349] In some of our patients with heterozygous *CNTNAP2* or *NRXN1* aberrations (**chapter 5)** additional unknown CNVs were observed by molecular karyotyping. However, an interpretation regarding a possible contributory effect of these CNVs is currently not yet possible.

The advances in NGS during the next years will hopefully help to understand how many cases of ID can indeed be explained by monogenic causes and how many will be due to a combination of multiple *de novo* and/or inherited defects.

7.4. From single genes to functional modules and back

Intellectual disability is clinically and genetically highly heterogeneous. The underlying genes are only incompletely identified so far, and even for the majority of known ID genes their individual function as well as their role in larger networks and pathways is only poorly understood.

Linking ID genes in common molecular or functional pathways indicates a general regulatory role of these networks in cognitive function and dysfunction[2,318-321] and supports the idea that similar phenotypes are caused by defects in functionally related genes.[178] By linking CNTNAP2 and NRXN1 via a common synaptic target, this thesis **(chapter 4)** contributed to this theme. To cope with the arising challenges resulting from NGS technologies and to further identify and delineate the complex networks of cognitive function and dysfunction, a systematic, large-scale and collaborative interdisciplinary approach is required.

Though several attempts were undertaken to create catalogs of ID genes, none of them was either highly reliable or complete.[96,257,325] We therefore established such an ID gene collection and introduced it into a database that is supplemented with many other datasets **(chapter 6)**. We furthermore established a classification system of ID genes, based on the clinical manifestation and severity of the associated disorders. This resulted in nine individual classes that could be summarized in six overlapping higher-order groups (SWSM: syndromic with structural malformations, SWOSM, syndromic without structural malformations, NS: non-syndromic; CS: classic severe, CM: classic mild to moderate or variable, NC: non-classic). First analyses showed several interesting enrichment and expression patterns **(chapter 6)**:

1) Genes from the postsynaptic density of the human neocortex[322] are significantly enriched in the non-syndromic group (NS). This goes in hand with significantly increased expression levels for NS genes in fetal brain and several brain regions. Enrichment of X-linked inheritance in the NS group might indicate a high proportion of ID genes with synaptic function on the X-chromosome, as was also discussed previously.[329]

2) The highest enrichment of haploinsufficiency (HI) genes is found in the SWSM group. This goes in hand with significant enrichment of autosomal dominant inheritance (in terms of *de novo* mutations) in the SWSM group. These observations indicate the role of HI genes in disorders that affect the development of the brain and many other organs and systems.

3) Autosomal-recessive inheritance is significantly enriched in the SWOSM group while it is significantly depleted in the SWSM and the NS groups. This indicates that mutations in autosomal recessive genes cause syndromic ID but without structural malformations which might be true for most of the metabolic genes.

4) CS genes show increased expression in various brain regions, while CM genes show no significant expression pattern, and NC genes shows a heterogeneous pattern for organs, brain tissues and blood and immune cells. Thus, the high and specific expression of CS genes in brain is correlating with the severity of ID.

These analyses are only our first steps towards a comprehensive systems biology-like approach. They already demonstrated that both molecular as well as phenotypic "grouping" can be used to detect patterns. These might be helpful both for ID disease gene prediction as well as for prediction of protein function and modularity.

7.5. Back to the future

The growing impact and progression in NGS technology and data interpretation will rapidly increase the identification of new disease genes and of disease causing mutations in known genes in the next few years. This might reduce the importance of good clinical characterization prior to molecular testing, but it will raise the need for good clinical phenotyping afterwards. As NGS will result in the detection of various mutations in various genes, a good interaction between clinical and molecular findings will be the prerequisite to accurately delineate genotype-phenotype correlations.

Only recently, approaches like multiplex targeting sequences (MIB) were developed for cost-efficient resequencing of candidate genes in large cohorts of patients with ID or epilepsy.[360,361] Nevertheless, due to the extreme heterogeneity in ID, it has to be assumed that not for all patients or families with a mutation in a convincing candidate gene confirmation for its pathogenicity can be obtained by detecting other patients or families with mutations in the same gene easily. This will be particularly relevant for missense variants, for which a deleterious effect is less obvious than for truncating mutations. For some cases, functional analyses will be and remain necessary to support a pathogenic relevance of particular genetic defects.

It can also be safely assumed that not all ID disorders will be due to single defects in a single gene, but that there will be certain proportion of oligogenic or even multifactorial causes of ID, particularly in the milder spectrum. Large clinical variability in some monogenic ID disorders - sometimes even resulting from an identical mutation - is assumed to result from additional, modifying genetic, epigenetic and environmental factors. To delineate at least the former will require large bioinformatics and statistical approaches on a massive amount of data. It will also require a thorough functional characterization of genes and variants in order to map them to common pathways and complexes and to explain their influence on each other.

Understanding the molecular and functional basis of ID and to connect single genes and proteins in functional networks and pathways is the prerequisite to get back to clinics and establish therapeutic approaches. This is one of the big challenges for the years and decades to come.

Taken together, bringing together clinics, molecular diagnostics, basic research, bioinformatics and pharmacology for the ultimate goal of unraveling and treating ID is both a challenge and a promise for the future of human genetics and medical genomics.

Summary

Intellectual disability (ID) affects 2-3% of the population and therefore has a major social and economic impact on both families and society. The genetic causes for ID are very heterogeneous and so far, the underlying cause can only be identified in about half of the affected individuals. During the last years the development of first molecular karyotyping in order to detect small chromosomal aberrations and then Next-Generation-Sequencing to detect sequence variants provided potent new technologies to identify known and novel genetic causes of ID. Diagnosing the underlying genetic defect in a particular family helps to provide proper counseling regarding prognosis, potential complications, comorbidities, and recurrence risks on one hand. On the other hand, identification of novel genes and connecting them in functional networks provides the basis for the long-term development of therapeutic approaches. For this, a detailed understanding of the molecular and functional basis of ID is necessary as well as connecting single genes and proteins in functional networks and pathways. This requires a tight collaboration of different disciplines such as clinics, molecular diagnostics, basic research, and bioinformatics.

The aim of this thesis was to contribute to the genetic, clinical and functional characterization of ID disorders, with a particular focus on Pitt-Hopkins (PTHS) and Pitt-Hopkins-like syndromes.

Chapter 1 provides an introduction on intellectual disability. It discusses the genetic and clinical heterogeneity, the underlying molecular defects, methods for disease gene identification, and functional testing with different approaches.

Chapter 2 describes the detection of a *TCF4* deletion by molecular karyotyping in a patient with Pitt-Hopkins syndrome. This disorder is characterized by severe ID, breathing anomalies and typical facial dysmorphism. Subsequent identification of mutations in this gene in further patients with a similar clinical phenotype identified *TCF4* haploinsufficiency as the underlying cause for PTHS and contributed to the delineation of the distinct, but until then under-recognized, phenotype. Impaired transcriptional interaction with ASCL1 from the ASCL1-PHOX2B-Ret pathway upon TCF4 mutations was demonstrated with a transcriptional reporter assay and might explain some of the characteristic PTHS symptoms.

In **chapter 3**, a larger number of patients with PTHS and TCF4 mutations were collected in order to further delineate the clinical and genetic characteristics of this syndrome. This helped to increase the awareness of clinicians towards this disorder and to establish it as an important differential diagnosis to Rett- and Angelman syndromes, widely known and intensely studied severe ID disorders.

As only in a subset of patients referred with suspected PTHS mutations in *TCF4* could be identified, c**hapter 4** describes the continuing screening for small aberrations with molecular karyotyping in the remaining patients. This revealed a homozygous deletion within *CNTNAP2* in a pair of siblings, and a further sporadic patient with a compound heterozygous deletion and mutation in this gene. Additionally, a compound heterozygous deletion and mutation in the *NRXN1* gene was identified in a sporadic patient with a similar phenotype. *Drosophila melanogaster* as a model organism allowed to demonstrate a possible interaction between Nrx-IV (CNTNAP2) and Nrx-I (NRXN1) by converging on a common target, the presynaptic protein bruchpilot.

Chapter 5 describes how heterozygous defects in *CTNAP2* and *NRXN1* but no second mutation were found in patients with severe ID and one of their healthy parents, respectively. This expanded the clinical spectrum associated with heterozygous defects in any of the two genes to the severe end by including severe ID. This observation suggested the existence of additionally contributing genetic factors, thus pointing to a possible di- or oligogenic cause for the severe ID in these patients

The work described in **chapter 6** aims at a more global understanding of the molecular pathology of ID disorders by establishing a systematic inventory of all ID-related genes and their associated phenotypes. By creating a clinical classification system based on the manifestation of ID as well as on accompanying phenotypic features this study attempted to establish correlations between clinical and molecular/functional aspects.

A general discussion of the findings in this thesis and their implications in the broader field of ID genetics as well as future perspectives are provided in **chapter 7**. Understanding the molecular and functional basis of ID by connecting single genes and proteins in functional networks and pathways is the prerequisite to begin to unravel and treat ID in the coming years.

References

1. Martin, G. 'Valuing people'--a new strategy for learning disability for the 21st century: how may it impinge on primary care? *Br J Gen Pract* **51**, 788-90 (2001).
2. Ropers, H.H. Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* **11**, 161-87 (2010).
3. van Bokhoven, H. Genetic and epigenetic networks in intellectual disabilities. *Annu Rev Genet* **45**, 81-104 (2011).
4. Vaillend, C., Poirier, R. & Laroche, S. Genes, plasticity and mental retardation. *Behav Brain Res* **192**, 88-105 (2008).
5. Lejeune, J., Turpin, R. & Gautier, M. [Mongolism; a chromosomal disease (trisomy)]. *Bull Acad Natl Med* **143**, 256-65 (1959).
6. Lejeune, J., Gautier, M. & Turpin, R. [Study of somatic chromosomes from 9 mongoloid children]. *C R Hebd Seances Acad Sci* **248**, 1721-2 (1959).
7. Pinkel, D., Landegent, J., Collins, C. *et al.* Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc Natl Acad Sci U S A* **85**, 9138-42 (1988).
8. Schouten, J.P., McElgunn, C.J., Waaijer, R. *et al.* Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* **30**, e57 (2002).
9. Pinkel, D., Segraves, R., Sudar, D. *et al.* High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* **20**, 207-11 (1998).
10. Rauch, A., Ruschendorf, F., Huang, J. *et al.* Molecular karyotyping using an SNP array for genomewide genotyping. *J Med Genet* **41**, 916-22 (2004).
11. Vissers, L.E., de Vries, B.B., Osoegawa, K. *et al.* Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet* **73**, 1261-70 (2003).
12. Kremer, E.J., Pritchard, M., Lynch, M. *et al.* Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* **252**, 1711-4 (1991).
13. Ng, S.B., Turner, E.H., Robertson, P.D. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272-6 (2009).
14. Rauch, A., Hoyer, J., Guth, S. *et al.* Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet A* **140**, 2063-74 (2006).
15. Koolen, D.A., Pfundt, R., de Leeuw, N. *et al.* Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. *Hum Mutat* **30**, 283-92 (2009).
16. Sagoo, G.S., Butterworth, A.S., Sanderson, S. *et al.* Array CGH in patients with learning disability (mental retardation) and congenital anomalies: updated systematic review and meta-analysis of 19 studies and 13,926 subjects. *Genet Med* **11**, 139-46 (2009).
17. Zollino, M., Di Stefano, C., Zampino, G. *et al.* Genotype-phenotype correlations and clinical diagnostic criteria in Wolf-Hirschhorn syndrome. *Am J Med Genet* **94**, 254-61 (2000).
18. Shapira, S.K., McCaskill, C., Northrup, H. *et al.* Chromosome 1p36 deletions: the clinical phenotype and molecular characterization of a common newly delineated syndrome. *Am J Hum Genet* **61**, 642-50 (1997).
19. Shaw, C.J. & Lupski, J.R. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet* **13 Spec No 1**, R57-64 (2004).
20. Stankiewicz, P. & Lupski, J.R. Structural variation in the human genome and its role in disease. *Annu Rev Med* **61**, 437-55 (2010).
21. Driscoll, D.A., Budarf, M.L. & Emanuel, B.S. A genetic etiology for DiGeorge syndrome: consistent deletions and microdeletions of 22q11. *Am J Hum Genet* **50**, 924-33 (1992).
22. Bayes, M., Magano, L.F., Rivera, N., Flores, R. & Perez Jurado, L.A. Mutational mechanisms of Williams-Beuren syndrome deletions. *Am J Hum Genet* **73**, 131-51 (2003).
23. Koolen, D.A., Vissers, L.E., Pfundt, R. *et al.* A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* **38**, 999-1001 (2006).
24. Sharp, A.J., Hansen, S., Selzer, R.R. *et al.* Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* **38**, 1038-42 (2006).
25. Shaw-Smith, C., Pittman, A.M., Willatt, L. *et al.* Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet* **38**, 1032-7 (2006).
26. Mefford, H.C., Cooper, G.M., Zerr, T. *et al.* A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res* **19**, 1579-85 (2009).

27. Osborne, L.R., Martindale, D., Scherer, S.W. *et al.* Identification of genes from a 500-kb region at 7q11.23 that is commonly deleted in Williams syndrome patients. *Genomics* **36**, 328-36 (1996).
28. Baldini, A. Dissecting contiguous gene defects: TBX1. *Curr Opin Genet Dev* **15**, 279-84 (2005).
29. Kleefstra, T., Brunner, H.G., Amiel, J. *et al.* Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am J Hum Genet* **79**, 370-7 (2006).
30. Gong, W., Gottlieb, S., Collins, J. *et al.* Mutation analysis of TBX1 in non-deleted patients with features of DGS/VCFS or isolated cardiovascular defects. *J Med Genet* **38**, E45 (2001).
31. Zweier, C., Sticht, H., Aydin-Yaylagul, I., Campbell, C.E. & Rauch, A. Human TBX1 missense mutations cause gain of function resulting in the same phenotype as 22q11.2 deletions. *Am J Hum Genet* **80**, 510-7 (2007).
32. Leana-Cox, J., Pangkanon, S., Eanet, K.R., Curtin, M.S. & Wulfsberg, E.A. Familial DiGeorge/velocardiofacial syndrome with deletions of chromosome area 22q11.2: report of five families with a review of the literature. *Am J Med Genet* **65**, 309-16 (1996).
33. Bijlsma, E.K., Gijsbers, A.C., Schuurs-Hoeijmakers, J.H. *et al.* Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur J Med Genet* **52**, 77-87 (2009).
34. de Ligt, J., Willemsen, M.H., van Bon, B.W. *et al.* Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *N Engl J Med* (2012).
35. Rauch, A., Wieczorek, D., Graf, E. *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* (2012).
36. Vissers, L.E., van Ravenswaaij, C.M., Admiraal, R. *et al.* Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* **36**, 955-7 (2004).
37. Verkerk, A.J., Pieretti, M., Sutcliffe, J.S. *et al.* Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905-14 (1991).
38. de Vries, B.B., van den Ouweland, A.M., Mohkamsing, S. *et al.* Screening and diagnosis for the fragile X syndrome among the mentally retarded: an epidemiological and psychological survey. Collaborative Fragile X Study Group. *Am J Hum Genet* **61**, 660-7 (1997).
39. Billuart, P., Bienvenu, T., Ronce, N. *et al.* Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental retardation. *Nature* **392**, 923-6 (1998).
40. Tarpey, P.S., Smith, R., Pleasance, E. *et al.* A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* **41**, 535-43 (2009).
41. Ropers, H.H. & Hamel, B.C. X-linked mental retardation. *Nat Rev Genet* **6**, 46-57 (2005).
42. Vissers, L.E., de Ligt, J., Gilissen, C. *et al.* A de novo paradigm for mental retardation. *Nat Genet* **42**, 1109-12 (2010).
43. Zenker, M., Voss, E. & Reis, A. Mild variable Noonan syndrome in a family with a novel PTPN11 mutation. *Eur J Med Genet* **50**, 43-7 (2007).
44. Vafai, S.B. & Mootha, V.K. Mitochondrial disorders as windows into an ancient organelle. *Nature* **491**, 374-83 (2012).
45. Girirajan, S., Rosenfeld, J.A., Cooper, G.M. *et al.* A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* **42**, 203-9 (2010).
46. Leblond, C.S., Heinrich, J., Delorme, R. *et al.* Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet* **8**, e1002521 (2012).
47. O'Roak, B.J., Vives, L., Girirajan, S. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246-50 (2012).
48. Katsanis, N. The oligogenic properties of Bardet-Biedl syndrome. *Hum Mol Genet* **13 Spec No 1**, R65-71 (2004).
49. Chelly, J., Khelfaoui, M., Francis, F., Cherif, B. & Bienvenu, T. Genetics and pathophysiology of mental retardation. *Eur J Hum Genet* **14**, 701-13 (2006).
50. Klopocki, E. & Mundlos, S. Copy-number variations, noncoding sequences, and human phenotypes. *Annu Rev Genomics Hum Genet* **12**, 53-72 (2011).
51. Hobson, G.M., Huang, Z., Sperle, K. *et al.* A PLP splicing abnormality is associated with an unusual presentation of PMD. *Ann Neurol* **52**, 477-88 (2002).
52. Huang, L., Jolly, L.A., Willis-Owen, S. *et al.* A noncoding, regulatory mutation implicates HCFC1 in nonsyndromic intellectual disability. *Am J Hum Genet* **91**, 694-702 (2012).
53. Botstein, D. & Risch, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* **33 Suppl**, 228-37 (2003).

54. Kerem, B., Rommens, J.M., Buchanan, J.A. *et al.* Identification of the cystic fibrosis gene: genetic analysis. *Science* **245**, 1073-80 (1989).
55. Vance, J.M., Pericak-Vance, M.A., Yamaoka, L.H. *et al.* Genetic linkage mapping of chromosome 17 markers and neurofibromatosis type I. *Am J Hum Genet* **44**, 25-9 (1989).
56. Wallace, M.R., Marchuk, D.A., Andersen, L.B. *et al.* Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients. *Science* **249**, 181-6 (1990).
57. Kolehmainen, J., Black, G.C., Saarinen, A. *et al.* Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *Am J Hum Genet* **72**, 1359-69 (2003).
58. Lander, E.S. & Botstein, D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* **236**, 1567-70 (1987).
59. Bittles, A. Consanguinity and its relevance to clinical genetics. *Clin Genet* **60**, 89-98 (2001).
60. Ropers, H.H. New perspectives for the elucidation of genetic disorders. *Am J Hum Genet* **81**, 199-207 (2007).
61. Abou Jamra, R., Wohlfart, S., Zweier, M. *et al.* Homozygosity mapping in 64 Syrian consanguineous families with non-specific intellectual disability reveals 11 novel loci and high heterogeneity. *Eur J Hum Genet* **19**, 1161-6 (2011).
62. Najmabadi, H., Motazacker, M.M., Garshasbi, M. *et al.* Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum Genet* **121**, 43-8 (2007).
63. Najmabadi, H., Hu, H., Garshasbi, M. *et al.* Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* **478**, 57-63 (2011).
64. Warburton, D. De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am J Hum Genet* **49**, 995-1013 (1991).
65. Wirth, J., Nothwang, H.G., van der Maarel, S. *et al.* Systematic characterisation of disease associated balanced chromosome rearrangements by FISH: cytogenetically and genetically anchored YACs identify microdeletions and candidate regions for mental retardation genes. *J Med Genet* **36**, 271-8 (1999).
66. Yamada, K., Yamada, Y., Nomura, N. *et al.* Nonsense and frameshift mutations in ZFHX1B, encoding Smad-interacting protein 1, cause a complex developmental disorder with a great variety of clinical features. *Am J Hum Genet* **69**, 1178-85 (2001).
67. Endele, S., Rosenberger, G., Geider, K. *et al.* Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* **42**, 1021-6 (2010).
68. Kalscheuer, V.M., Tao, J., Donnelly, A. *et al.* Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *Am J Hum Genet* **72**, 1401-11 (2003).
69. Tao, J., Van Esch, H., Hagedorn-Greiwe, M. *et al.* Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5/STK9) gene are associated with severe neurodevelopmental retardation. *Am J Hum Genet* **75**, 1149-54 (2004).
70. de Vries, B.B., Pfundt, R., Leisink, M. *et al.* Diagnostic genome profiling in mental retardation. *Am J Hum Genet* **77**, 606-16 (2005).
71. Hoyer, J., Dreweke, A., Becker, C. *et al.* Molecular karyotyping in patients with mental retardation using 100K single-nucleotide polymorphism arrays. *J Med Genet* **44**, 629-36 (2007).
72. Slavotinek, A.M. Novel microdeletion syndromes detected by chromosome microarrays. *Hum Genet* **124**, 1-17 (2008).
73. Lesnik Oberstein, S.A., Kriek, M., White, S.J. *et al.* Peters Plus syndrome is caused by mutations in B3GALTL, a putative glycosyltransferase. *Am J Hum Genet* **79**, 562-6 (2006).
74. Saitsu, H., Kato, M., Mizuguchi, T. *et al.* De novo mutations in the gene encoding STXBP1 (MUNC18-1) cause early infantile epileptic encephalopathy. *Nat Genet* **40**, 782-8 (2008).
75. Zweier, M., Gregor, A., Zweier, C. *et al.* Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum Mutat* **31**, 722-33 (2010).
76. Hoyer, J., Ekici, A.B., Endele, S. *et al.* Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am J Hum Genet* **90**, 565-72 (2012).
77. Hamdan, F.F., Gauthier, J., Spiegelman, D. *et al.* Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. *N Engl J Med* **360**, 599-605 (2009).
78. Hamdan, F.F., Gauthier, J., Araki, Y. *et al.* Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am J Hum Genet* **88**, 306-16 (2011).

79. Arts, H.H., Doherty, D., van Beersum, S.E. *et al.* Mutations in the gene encoding the basal body protein RPGRIP1L, a nephrocystin-4 interactor, cause Joubert syndrome. *Nat Genet* **39**, 882-8 (2007).

80. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat Biotechnol* **26**, 1135-45 (2008).

81. Kahrizi, K., Hu, C.H., Garshasbi, M. *et al.* Next generation sequencing in a family with autosomal recessive Kahrizi syndrome (OMIM 612713) reveals a homozygous frameshift mutation in SRD5A3. *Eur J Hum Genet* **19**, 115-7 (2011).

82. Abou Jamra, R., Philippe, O., Raas-Rothschild, A. *et al.* Adaptor protein complex 4 deficiency causes severe autosomal-recessive intellectual disability, progressive spastic paraplegia, shy character, and short stature. *Am J Hum Genet* **88**, 788-95 (2011).

83. Ng, S.B., Bigham, A.W., Buckingham, K.J. *et al.* Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* **42**, 790-3 (2010).

84. Hoischen, A., van Bon, B.W., Gilissen, C. *et al.* De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet* **42**, 483-5 (2010).

85. Deardorff, M.A., Clark, D.M. & Krantz, I.D. Cornelia de Lange Syndrome. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

86. Dobyns, W.B. & Das, S. LIS1-Associated Lissencephaly/Subcortical Band Heterotopia. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

87. Morris, C.A. Williams Syndrome. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

88. Adam, M.P., Bean, L.J.H. & Miller, V.R. Mowat-Wilson Syndrome. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

89. Christodoulou, J. & Ho, G. MECP2-Related Disorders. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

90. Steinberg, S.J., Moser, A.B. & Raymond, G.V. X-Linked Adrenoleukodystrophy. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

91. Garcia-Cazorla, A., Wolf, N.I., Serrano, M. *et al.* Mental retardation and inborn errors of metabolism. *J Inherit Metab Dis* **32**, 597-608 (2009).

92. Friedman, J.M. Neurofibromatosis 1. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

93. Fisch, G.S. Nosology and epidemiology in autism: classification counts. *Am J Med Genet C Semin Med Genet* **160C**, 91-103 (2012).

94. Oeseburg, B., Dijkstra, G.J., Groothoff, J.W., Reijneveld, S.A. & Jansen, D.E. Prevalence of chronic health conditions in children with intellectual disability: a systematic literature review. *Intellect Dev Disabil* **49**, 59-85 (2011).

95. Schwartz, C.E. & Neri, G. Autism and intellectual disability: two sides of the same coin. *Am J Med Genet C Semin Med Genet* **160C**, 89-90 (2012).

96. Kou, Y., Betancur, C., Xu, H., Buxbaum, J.D. & Ma'ayan, A. Network- and attribute-based classifiers can prioritize genes and pathways for autism spectrum disorders and intellectual disability. *Am J Med Genet C Semin Med Genet* **160C**, 130-42 (2012).

97. Lana-Elola, E., Watson-Scales, S.D., Fisher, E.M. & Tybulewicz, V.L. Down syndrome: searching for the genetic culprits. *Dis Model Mech* **4**, 586-95 (2011).

98. van Bon, B.W., Hoischen, A., Hehir-Kwa, J. *et al.* Intragenic deletion in DYRK1A leads to mental retardation and primary microcephaly. *Clin Genet* **79**, 296-9 (2011).

99. Baldini, A. The 22q11.2 deletion syndrome: a gene dosage perspective. *ScientificWorldJournal* **6**, 1881-7 (2006).

100. Smith, A.C., McGavran, L., Robinson, J. *et al.* Interstitial deletion of (17)(p11.2p11.2) in nine patients. *Am J Med Genet* **24**, 393-414 (1986).

101. Potocki, L., Bi, W., Treadwell-Deering, D. *et al.* Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. *Am J Hum Genet* **80**, 633-49 (2007).

102. Ramocki, M.B. & Zoghbi, H.Y. Failure of neuronal homeostasis results in common neuropsychiatric phenotypes. *Nature* **455**, 912-8 (2008).

103. Simpson, M.A., Deshpande, C., Dafou, D. *et al.* De novo mutations of the gene encoding the histone acetyltransferase KAT6B cause Genitopatellar syndrome. *Am J Hum Genet* **90**, 290-4 (2012).

104. Campeau, P.M., Kim, J.C., Lu, J.T. *et al.* Mutations in KAT6B, encoding a histone acetyltransferase, cause Genitopatellar syndrome. *Am J Hum Genet* **90**, 282-9 (2012).

105. Clayton-Smith, J., O'Sullivan, J., Daly, S. *et al.* Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am J Hum Genet* **89**, 675-81 (2011).

106. Campeau, P.M., Lu, J.T., Dawson, B.C. *et al.* The KAT6B-related disorders genitopatellar syndrome and Ohdo/SBBYS syndrome have distinct clinical features reflecting distinct molecular mechanisms. *Hum Mutat* **33**, 1520-5 (2012).

107. Frischmeyer, P.A. & Dietz, H.C. Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet* **8**, 1893-900 (1999).

108. Ghosh, R.P., Horowitz-Scherer, R.A., Nikitina, T., Gierasch, L.M. & Woodcock, C.L. Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions. *J Biol Chem* **283**, 20523-34 (2008).

109. Zenker, M. Genetic and pathogenetic aspects of Noonan syndrome and related disorders. *Horm Res* **72 Suppl 2**, 57-63 (2009).

110. Marseglia, G., Scordo, M.R., Pescucci, C. *et al.* 372 kb microdeletion in 18q12.3 causing SETBP1 haploinsufficiency associated with mild mental retardation and expressive speech impairment. *Eur J Med Genet* **55**, 216-21 (2012).

111. Vinkhuyzen, A.A., van der Sluis, S. & Posthuma, D. Life events moderate variation in cognitive ability (g) in adults. *Mol Psychiatry* **16**, 4-6 (2011).

112. Rizzi, T.S., Beunders, G., Rizzu, P. *et al.* Supporting the generalist genes hypothesis for intellectual ability/disability: the case of SNAP25. *Genes Brain Behav* **11**, 767-71 (2012).

113. Haworth, C.M. & Plomin, R. Quantitative genetics in the era of molecular genetics: learning abilities and disabilities as an example. *J Am Acad Child Adolesc Psychiatry* **49**, 783-93 (2010).

114. Vissers, L.E., de Vries, B.B. & Veltman, J.A. Genomic microarrays in mental retardation: from copy number variation to gene, from research to diagnosis. *J Med Genet* **47**, 289-97 (2010).

115. Kearney, J.A., Plummer, N.W., Smith, M.R. *et al.* A gain-of-function mutation in the sodium channel gene Scn2a results in seizures and behavioral abnormalities. *Neuroscience* **102**, 307-17 (2001).

116. Sanders, S.J., Murtha, M.T., Gupta, A.R. *et al.* De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237-41 (2012).

117. Van Houdt, J.K., Nowakowska, B.A., Sousa, S.B. *et al.* Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome. *Nat Genet* **44**, 445-9, S1 (2012).

118. Wolff, D., Endele, S., Azzarello-Burri, S. *et al.* In-Frame Deletion and Missense Mutations of the C-Terminal Helicase Domain of SMARCA2 in Three Patients with Nicolaides-Baraitser Syndrome. *Molecular Syndromology* **2**, 237-244 (2011).

119. Brook, J.D., McCurrach, M.E., Harley, H.G. *et al.* Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**, 799-808 (1992).

120. Ranum, L.P. & Day, J.W. Myotonic dystrophy: RNA pathogenesis comes into focus. *Am J Hum Genet* **74**, 793-804 (2004).

121. Taneja, K.L., McCurrach, M., Schalling, M., Housman, D. & Singer, R.H. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* **128**, 995-1002 (1995).

122. Nicholls, R.D. The impact of genomic imprinting for neurobehavioral and developmental disorders. *J Clin Invest* **105**, 413-8 (2000).

123. Kahler, S.G. & Fahey, M.C. Metabolic disorders and mental retardation. *Am J Med Genet C Semin Med Genet* **117C**, 31-41 (2003).

124. Yaghootfam, A., Gieselmann, V. & Eckhardt, M. Delay of myelin formation in arylsulphatase A-deficient mice. *Eur J Neurosci* **21**, 711-20 (2005).

125. Huttenlocher, P.R. The neuropathology of phenylketonuria: human and animal studies. *Eur J Pediatr* **159 Suppl 2**, S102-6 (2000).

126. Jackson, A.P., Eastwood, H., Bell, S.M. *et al.* Identification of microcephalin, a protein implicated in determining the size of the human brain. *Am J Hum Genet* **71**, 136-42 (2002).

127. Bond, J., Roberts, E., Mochida, G.H. *et al.* ASPM is a major determinant of cerebral cortical size. *Nat Genet* **32**, 316-20 (2002).

128. Bond, J., Roberts, E., Springell, K. *et al.* A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size. *Nat Genet* **37**, 353-5 (2005).

129. Verrotti, A., Spalice, A., Ursitti, F. *et al.* New trends in neuronal migration disorders. *Eur J Paediatr Neurol* **14**, 1-12 (2010).

130. Grant, S.G. Synaptopathies: diseases of the synaptome. *Curr Opin Neurobiol* **22**, 522-9 (2012).

131. D'Adamo, P., Menegon, A., Lo Nigro, C. *et al.* Mutations in GDI1 are responsible for X-linked non-specific mental retardation. *Nat Genet* **19**, 134-9 (1998).

132. Giannandrea, M., Bianchi, V., Mignogna, M.L. *et al.* Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. *Am J Hum Genet* **86**, 185-95 (2010).

133. Najm, J., Horn, D., Wimplinger, I. *et al.* Mutations of CASK cause an X-linked brain malformation phenotype with microcephaly and hypoplasia of the brainstem and cerebellum. *Nat Genet* **40**, 1065-7 (2008).

134. Wu, Y., Arai, A.C., Rumbaugh, G. *et al.* Mutations in ionotropic AMPA receptor 3 alter channel properties and are associated with moderate cognitive impairment in humans. *Proc Natl Acad Sci U S A* **104**, 18163-8 (2007).

135. Berkel, S., Marshall, C.R., Weiss, B. *et al.* Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation. *Nat Genet* **42**, 489-91 (2010).

136. Sato, D., Lionel, A.C., Leblond, C.S. *et al.* SHANK1 Deletions in Males with Autism Spectrum Disorder. *Am J Hum Genet* **90**, 879-87 (2012).

137. Moessner, R., Marshall, C.R., Sutcliffe, J.S. *et al.* Contribution of SHANK3 mutations to autism spectrum disorder. *Am J Hum Genet* **81**, 1289-97 (2007).

138. Jin, P., Zarnescu, D.C., Ceman, S. *et al.* Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. *Nat Neurosci* **7**, 113-7 (2004).

139. Matsuura, T., Sutcliffe, J.S., Fang, P. *et al.* De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome. *Nat Genet* **15**, 74-7 (1997).

140. Greer, P.L., Hanayama, R., Bloodgood, B.L. *et al.* The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. *Cell* **140**, 704-16 (2010).

141. Newey, S.E., Velamoor, V., Govek, E.E. & Van Aelst, L. Rho GTPases, dendritic structure, and mental retardation. *J Neurobiol* **64**, 58-74 (2005).

142. Allen, K.M., Gleeson, J.G., Bagrodia, S. *et al.* PAK3 mutation in nonsyndromic X-linked mental retardation. *Nat Genet* **20**, 25-30 (1998).

143. Kutsche, K., Yntema, H., Brandt, A. *et al.* Mutations in ARHGEF6, encoding a guanine nucleotide exchange factor for Rho GTPases, in patients with X-linked mental retardation. *Nat Genet* **26**, 247-50 (2000).

144. Pasteris, N.G., Cadle, A., Logie, L.J. *et al.* Isolation and characterization of the faciogenital dysplasia (Aarskog-Scott syndrome) gene: a putative Rho/Rac guanine nucleotide exchange factor. *Cell* **79**, 669-78 (1994).

145. Krab, L.C., Goorden, S.M. & Elgersma, Y. Oncogenes on my mind: ERK and MTOR signaling in cognitive diseases. *Trends Genet* **24**, 498-510 (2008).

146. Kratz, C.P., Schubbert, S., Bollag, G. *et al.* Germline mutations in components of the Ras signaling pathway in Noonan syndrome and related disorders. *Cell Cycle* **5**, 1607-11 (2006).

147. Tartaglia, M. & Gelb, B.D. Disorders of dysregulated signal traffic through the RAS-MAPK pathway: phenotypic spectrum and molecular mechanisms. *Ann N Y Acad Sci* **1214**, 99-121 (2010).

148. Cesarini, L., Alfieri, P., Pantaleoni, F. *et al.* Cognitive profile of disorders associated with dysregulation of the RAS/MAPK signaling cascade. *Am J Med Genet A* **149A**, 140-6 (2009).

149. Alberini, C.M. Transcription factors in long-term memory and synaptic plasticity. *Physiol Rev* **89**, 121-45 (2009).

150. Wakamatsu, N., Yamada, Y., Yamada, K. *et al.* Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nat Genet* **27**, 369-70 (2001).

151. Cacheux, V., Dastot-Le Moal, F., Kaariainen, H. *et al.* Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease. *Hum Mol Genet* **10**, 1503-10 (2001).

152. Le Meur, N., Holder-Espinasse, M., Jaillard, S. *et al.* MEF2C haploinsufficiency caused by either microdeletion of the 5q14.3 region or mutation is responsible for severe mental retardation with stereotypic movements, epilepsy and/or cerebral malformations. *J Med Genet* **47**, 22-9 (2010).

153. Laumonnier, F., Ronce, N., Hamel, B.C. *et al.* Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet* **71**, 1450-5 (2002).

154. Touraine, R.L., Attie-Bitach, T., Manceau, E. *et al.* Neurological phenotype in Waardenburg syndrome type 4 correlates with novel SOX10 truncating mutations and expression in developing brain. *Am J Hum Genet* **66**, 1496-503 (2000).

155. van Bokhoven, H. & Kramer, J.M. Disruption of the epigenetic code: an emerging mechanism in mental retardation. *Neurobiol Dis* **39**, 3-12 (2010).

156. Jones, P.L., Veenstra, G.J., Wade, P.A. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* **19**, 187-91 (1998).

157. Amir, R.E., Van den Veyver, I.B., Wan, M. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23**, 185-8 (1999).

158. Petrij, F., Giles, R.H., Dauwerse, H.G. *et al.* Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **376**, 348-51 (1995).

159. Roelfsema, J.H., White, S.J., Ariyurek, Y. *et al.* Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet* **76**, 572-80 (2005).

160. Trivier, E., De Cesare, D., Jacquot, S. *et al.* Mutations in the kinase Rsk-2 associated with Coffin-Lowry syndrome. *Nature* **384**, 567-70 (1996).

161. Wolff, D., Endele, S., Azzarello-Burri, S. *et al.* In-Frame Deletion and Missense Mutations of the C-Terminal Helicase Domain of SMARCA2 in Three Patients with Nicolaides-Baraitser Syndrome. *Mol Syndromol* **2**, 237-244 (2012).

162. Santen, G.W., Aten, E., Sun, Y. *et al.* Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet* **44**, 379-80 (2012).

163. Tsurusaki, Y., Okamoto, N., Ohashi, H. *et al.* Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* **44**, 376-8 (2012).

164. Halgren, C., Kjaergaard, S., Bak, M. *et al.* Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clin Genet* **82**, 248-255 (2012).

165. Peterson, C.L. & Workman, J.L. Promoter targeting and chromatin remodeling by the SWI/SNF complex. *Curr Opin Genet Dev* **10**, 187-92 (2000).

166. Gibbons, R.J., Picketts, D.J., Villard, L. & Higgs, D.R. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell* **80**, 837-45 (1995).

167. Riviere, J.B., van Bon, B.W., Hoischen, A. *et al.* De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. *Nat Genet* **44**, 440-4, S1-2 (2012).

168. Mulley, J.C., Scheffer, I.E., Petrou, S. & Berkovic, S.F. Channelopathies as a genetic cause of epilepsy. *Curr Opin Neurol* **16**, 171-6 (2003).

169. Mantegazza, M., Gambardella, A., Rusconi, R. *et al.* Identification of an Nav1.1 sodium channel (SCN1A) loss-of-function mutation associated with familial simple febrile seizures. *Proc Natl Acad Sci U S A* **102**, 18177-82 (2005).

170. Claes, L., Del-Favero, J., Ceulemans, B. *et al.* De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. *Am J Hum Genet* **68**, 1327-32 (2001).

171. Dibbens, L.M., Tarpey, P.S., Hynes, K. *et al.* X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. *Nat Genet* **40**, 776-81 (2008).

172. Hildebrandt, F., Benzing, T. & Katsanis, N. Ciliopathies. *N Engl J Med* **364**, 1533-43 (2011).

173. Lee, J.H. & Gleeson, J.G. The role of primary cilia in neuronal function. *Neurobiol Dis* **38**, 167-72 (2010).

174. Higginbotham, H., Eom, T.Y., Mariani, L.E. *et al.* Arl13b in primary cilia regulates the migration and placement of interneurons in the developing cerebral cortex. *Dev Cell* **23**, 925-38 (2012).

175. Parisi, M. & Glass, I. Joubert Syndrome and Related Disorders. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

176. Waters, A.M. & Beales, P.L. Bardet-Biedl Syndrome. in *GeneReviews* (eds. Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. & Adam, M.P.) (Seattle (WA), 1993).

177. McKusick, V.A. On lumpers and splitters, or the nosology of genetic disease. *Perspect Biol Med* **12**, 298-312 (1969).

178. Oti, M. & Brunner, H.G. The modular nature of genetic diseases. *Clin Genet* **71**, 1-11 (2007).

179. Zenker, M. Clinical manifestations of mutations in RAS and related intracellular signal transduction factors. *Curr Opin Pediatr* **23**, 443-51 (2011).

180. Zaghloul, N.A. & Katsanis, N. Functional modules, mutational load and human genetic disease. *Trends Genet* **26**, 168-76 (2010).

181. Oti, M., Huynen, M.A. & Brunner, H.G. Phenome connections. *Trends Genet* **24**, 103-6 (2008).

182. Buxbaum, J.D., Betancur, C., Bozdagi, O. *et al.* Optimizing the phenotyping of rodent ASD models: enrichment analysis of mouse and human neurobiological phenotypes associated with high-risk autism genes identifies morphological, electrophysiological, neurological, and behavioral features. *Mol Autism* **3**, 1 (2012).

183. Tartaglia, M., Gelb, B.D. & Zenker, M. Noonan syndrome and clinically related disorders. *Best Pract Res Clin Endocrinol Metab* **25**, 161-79 (2011).

184. Strauss, K.A., Puffenberger, E.G., Huentelman, M.J. *et al.* Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *N.Engl.J.Med.* **354**, 1370-1377 (2006).

185. Scorza, C.A. & Cavalheiro, E.A. Animal models of intellectual disability: towards a translational approach. *Clinics (Sao Paulo)* **66 Suppl 1**, 55-63 (2011).

186. Biancotti, J.C., Narwani, K., Buehler, N. *et al.* Human embryonic stem cells as models for aneuploid chromosomal syndromes. *Stem Cells* **28**, 1530-40 (2010).

187. Cundiff, P.E. & Anderson, S.A. Impact of induced pluripotent stem cells on the study of central nervous system disease. *Curr Opin Genet Dev* **21**, 354-61 (2011).

188. Marchetto, M.C., Carromeu, C., Acab, A. *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143**, 527-39 (2010).

189. Amenduni, M., De Filippis, R., Cheung, A.Y. *et al.* iPS cells to model CDKL5-related disorders. *Eur J Hum Genet* **19**, 1246-55 (2011).

190. Branchi, I., Bichler, Z., Berger-Sweeney, J. & Ricceri, L. Animal models of mental retardation: from gene to cognitive function. *Neurosci Biobehav Rev* **27**, 141-53 (2003).

191. Gatto, C.L. & Broadie, K. Drosophila modeling of heritable neurodevelopmental disorders. *Curr Opin Neurobiol* **21**, 834-41 (2011).

192. Bolduc, F.V. & Tully, T. Fruit flies and intellectual disability. *Fly (Austin)* **3**, 91-104 (2009).

193. Restifo, L.L. Mental retardation genes in drosophila: New approaches to understanding and treating developmental brain disorders. *Ment Retard Dev Disabil Res Rev* **11**, 286-94 (2005).

194. Bellen, H.J., Tong, C. & Tsuda, H. 100 years of Drosophila research and its impact on vertebrate neuroscience: a history lesson for the future. *Nat Rev Neurosci* **11**, 514-22 (2010).

195. Mayford, M. & Kandel, E.R. Genetic approaches to memory storage. *Trends Genet* **15**, 463-70 (1999).

196. McBride, S.M., Choi, C.H., Wang, Y. *et al.* Pharmacological rescue of synaptic plasticity, courtship behavior, and mushroom body defects in a Drosophila model of fragile X syndrome. *Neuron* **45**, 753-64 (2005).

197. Hagerman, R., Lauterborn, J., Au, J. & Berry-Kravis, E. Fragile X syndrome and targeted treatment trials. *Results Probl Cell Differ* **54**, 297-335 (2012).

198. Krueger, D.D. & Bear, M.F. Toward fulfilling the promise of molecular medicine in fragile X syndrome. *Annu Rev Med* **62**, 411-29 (2011).

199. Brand, A.H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-15 (1993).

200. Dietzl, G., Chen, D., Schnorrer, F. *et al.* A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **448**, 151-6 (2007).

201. Hannon, G.J. RNA interference. *Nature* **418**, 244-51 (2002).

202. Gohring, I., Tagariello, A., Endele, S. *et al.* Disruption of ST5 is associated with mental retardation and multiple congenital anomalies. *J Med Genet* **47**, 91-8 (2010).

203. Colak, D., Al-Dhalaan, H., Nester, M. *et al.* Genomic and transcriptomic analyses distinguish classic Rett and Rett-like syndrome and reveals shared altered pathways. *Genomics* **97**, 19-28 (2011).

204. Antonell, A., Vilardell, M. & Perez Jurado, L.A. Transcriptome profile in Williams-Beuren syndrome lymphoblast cells reveals gene pathways implicated in glucose intolerance and visuospatial construction deficits. *Hum Genet* **128**, 27-37 (2010).

205. Kraft, M., Cirstea, I.C., Voss, A.K. *et al.* Disruption of the histone acetyltransferase MYST4 leads to a Noonan syndrome-like phenotype and hyperactivated MAPK signaling in humans and mice. *J Clin Invest* **121**, 3479-91 (2011).

206. Engelen, E., Akinci, U., Bryne, J.C. *et al.* Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nat Genet* **43**, 607-11 (2011).

207. Merienne, K., Jacquot, S., Trivier, E. *et al.* Rapid immunoblot and kinase assay tests for a syndromal form of X linked mental retardation: Coffin-Lowry syndrome. *J Med Genet* **35**, 890-4 (1998).

208. Kamiya, K., Kaneda, M., Sugawara, T. *et al.* A nonsense mutation of the sodium channel gene SCN2A in a patient with intractable epilepsy and mental decline. *J Neurosci* **24**, 2690-8 (2004).

209. Oliva, M., Berkovic, S.F. & Petrou, S. Sodium channels and the neurobiology of epilepsy. *Epilepsia* (2012).

210. Phippard, D. & Manning, A.M. Screening for inhibitors of transcription factors using luciferase reporter gene expression in transfected cells. *Methods Mol Biol* **225**, 19-23 (2003).

211. Dey, B., Thukral, S., Krishnan, S. *et al.* DNA-protein interactions: methods for detection and analysis. *Mol Cell Biochem* **365**, 279-99 (2012).

212. Shoubridge, C., Tan, M.H., Seiboth, G. & Gecz, J. ARX homeodomain mutations abolish DNA binding and lead to a loss of transcriptional repression. *Hum Mol Genet* **21**, 1639-47 (2012).

213. Kramer, J.M., Kochinke, K., Oortveld, M.A. *et al.* Epigenetic regulation of learning and memory by Drosophila EHMT/G9a. *PLoS Biol* **9**, e1000569 (2011).

214. Genin, E.C., Geillon, F., Gondcaille, C. *et al.* Substrate specificity overlap and interaction between adrenoleukodystrophy protein (ALDP/ABCD1) and adrenoleukodystrophy-related protein (ALDRP/ABCD2). *J Biol Chem* **286**, 8075-84 (2011).

215. Buschdorf, J.P. & Stratling, W.H. A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. *J Mol Med (Berl)* **82**, 135-43 (2004).

216. Bruckner, A., Polge, C., Lentze, N., Auerbach, D. & Schlattner, U. Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci* **10**, 2763-88 (2009).

217. Todd, M.A. & Picketts, D.J. PHF6 Interacts with the Nucleosome Remodeling and Deacetylation (NuRD) Complex. *J Proteome Res* **11**, 4326-37 (2012).

218. Bardoni, B., Schenck, A. & Mandel, J.L. A novel RNA-binding nuclear protein that interacts with the fragile X mental retardation (FMR1) protein. *Hum Mol Genet* **8**, 2557-66 (1999).

219. Kuo, T.Y., Hong, C.J., Chien, H.L. & Hsueh, Y.P. X-linked mental retardation gene CASK interacts with Bcl11A/CTIP1 and regulates axon branching and outgrowth. *J Neurosci Res* **88**, 2364-73 (2010).

220. Gorden, N.T., Arts, H.H., Parisi, M.A. *et al.* CC2D2A is mutated in Joubert syndrome and interacts with the ciliopathy-associated basal body protein CEP290. *Am J Hum Genet* **83**, 559-71 (2008).

221. Mizutani, R., Nakamura, K., Yokoyama, S. *et al.* Developmental expression of sorting nexin 3 in the mouse central nervous system. *Gene Expr Patterns* **11**, 33-40 (2011).

222. Tomancak, P., Beaton, A., Weiszmann, R. *et al.* Systematic determination of patterns of gene expression during Drosophila embryogenesis. *Genome Biol* **3**, RESEARCH0088 (2002).

223. Ramakers, G.J., Wolfer, D., Rosenberger, G. *et al.* Dysregulation of Rho GTPases in the alphaPix/Arhgef6 mouse model of X-linked intellectual disability is paralleled by impaired structural and synaptic plasticity and cognitive deficits. *Hum Mol Genet* **21**, 268-86 (2012).

224. Balemans, M.C., Huibers, M.M., Eikelenboom, N.W. *et al.* Reduced exploration, increased anxiety, and altered social behavior: Autistic-like features of euchromatin histone methyltransferase 1 heterozygous knockout mice. *Behav Brain Res* **208**, 47-55 (2010).

225. Melicharek, D.J., Ramirez, L.C., Singh, S., Thompson, R. & Marenda, D.R. Kismet/CHD7 regulates axon morphology, memory and locomotion in a Drosophila model of CHARGE syndrome. *Hum Mol Genet* **19**, 4253-64 (2010).

226. Moy, S.S. & Nadler, J.J. Advances in behavioral genetics: mouse models of autism. *Mol Psychiatry* **13**, 4-26 (2008).

227. Pitt, D. & Hopkins, I. A syndrome of mental retardation, wide mouth and intermittent overbreathing. *Aust Paediatr J* **14**, 182-4 (1978).

228. Orrico, A., Galli, L., Zappella, M. *et al.* Possible case of Pitt-Hopkins syndrome in sibs. *Am.J.Med.Genet.* **103**, 157-159 (2001).

229. Peippo, M.M., Simola, K.O., Valanne, L.K. *et al.* Pitt-Hopkins syndrome in two patients and further definition of the phenotype. *Clin.Dysmorphol.* **15**, 47-54 (2006).

230. Singh, H.A. Mental retardation, macrostomia and hyperpnoea syndrome. *J Paediatr Child Health* **29**, 156-7 (1993).

231. Van, B.I., Quartel, S. & Hennekam, R.C. Mental retardation, "coarse" face, and hyperbreathing: confirmation of the Pitt-Hopkins syndrome. *Am.J.Med.Genet.* **75**, 273-276 (1998).

232. Zweier, C., Thiel, C.T., Dufke, A. *et al.* Clinical and mutational spectrum of Mowat-Wilson syndrome. *Eur J Med Genet* **48**, 97-111 (2005).

233. Corneliussen, B., Thornell, A., Hallberg, B. & Grundstrom, T. Helix-loop-helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers. *J Virol* **65**, 6084-93 (1991).

234. Bergqvist, I., Eriksson, M., Saarikettu, J. *et al.* The basic helix-loop-helix transcription factor E2-2 is involved in T lymphocyte development. *Eur J Immunol* **30**, 2857-63 (2000).

235. Zhuang, Y., Cheng, P. & Weintraub, H. B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. *Mol Cell Biol* **16**, 2898-905 (1996).

236. Persson, P., Jogi, A., Grynfeld, A., Pahlman, S. & Axelson, H. HASH-1 and E2-2 are expressed in human neuroblastoma cells and form a functional complex. *Biochem Biophys Res Commun* **274**, 22-31 (2000).

237. de Pontual, L., Nepote, V., Attie-Bitach, T. *et al.* Noradrenergic neuronal development is impaired by mutation of the proneural HASH-1 gene in congenital central hypoventilation syndrome (Ondine's curse). *Hum Mol Genet* **12**, 3173-80 (2003).

238. Dauger, S., Guimiot, F., Renolleau, S. *et al.* MASH-1/RET pathway involvement in development of brain stem control of respiratory frequency in newborn mice. *Physiol Genomics* **7**, 149-57 (2001).

239. Petersson, K., Ivars, F. & Sigvardsson, M. The pT alpha promoter and enhancer are direct targets for transactivation by E box-binding proteins. *Eur J Immunol* **32**, 911-20 (2002).

240. Amiel, J. & Lyonnet, S. Hirschsprung disease, associated syndromes, and genetics: a review. *J Med Genet* **38**, 729-39 (2001).

241. Berry-Kravis, E.M., Zhou, L., Rand, C.M. & Weese-Mayer, D.E. Congenital central hypoventilation syndrome: PHOX2B mutations and phenotype. *Am J Respir Crit Care Med* **174**, 1139-44 (2006).

242. Liu, Y., Ray, S.K., Yang, X.Q., Luntz-Leybman, V. & Chiu, I.M. A splice variant of E2-2 basic helix-loop-helix protein represses the brain-specific fibroblast growth factor 1 promoter through the binding to an imperfect E-box. *J Biol Chem* **273**, 19269-76 (1998).

243. Gustavsson, P., Kimber, E., Wahlstrom, J. & Anneren, G. Monosomy 18q syndrome and atypical Rett syndrome in a girl with an interstitial deletion (18)(q21.1q22.3). *Am J Med Genet* **82**, 348-51 (1999).

244. Sanlaville, D., Lapierre, J.M., Turleau, C. *et al.* Molecular karyotyping in human constitutional cytogenetics. *Eur J Med Genet* **48**, 214-31 (2005).

245. Vermeesch, J.R., Melotte, C., Froyen, G. *et al.* Molecular karyotyping: array CGH quality criteria for constitutional genetic diagnosis. *J Histochem Cytochem* **53**, 413-22 (2005).

246. Dastot-Le Moal, F., Wilson, M., Mowat, D. *et al.* ZFHX1B mutations in patients with Mowat-Wilson syndrome. *Hum Mutat* (2007).

247. Amiel, J., Rio, M., de, P.L. *et al.* Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am.J.Hum.Genet.* **80**, 988-993 (2007).

248. Zweier, C., Peippo, M.M., Hoyer, J. *et al.* Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am.J.Hum.Genet.* **80**, 994-1001 (2007).

249. Andrieux, J., Lepretre, F., Cuisset, J.M. *et al.* Deletion 18q21.2q21.32 involving TCF4 in a boy diagnosed by CGH-array. *Eur.J.Med.Genet.* **51**, 172-177 (2008).

250. Brockschmidt, A., Todt, U., Ryu, S. *et al.* Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum Mol Genet* **16**, 1488-94 (2007).

251. Kageyama, R. & Nakanishi, S. Helix-loop-helix factors in growth and differentiation of the vertebrate nervous system. *Curr Opin Genet Dev* **7**, 659-65 (1997).

252. Soosaar, A., Chiaramello, A., Zuber, M.X. & Neuman, T. Expression of basic-helix-loop-helix transcription factor ME2 during brain development and in the regions of neuronal plasticity in the adult brain. *Brain Res Mol Brain Res* **25**, 176-80 (1994).

253. Linding, R., Jensen, L.J., Diella, F. *et al.* Protein disorder prediction: implications for structural proteomics. *Structure* **11**, 1453-9 (2003).

254. Puntervoll, P., Linding, R., Gemund, C. *et al.* ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res* **31**, 3625-30 (2003).

255. Conchillo-Sole, O., de Groot, N.S., Aviles, F.X. *et al.* AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics* **8**, 65 (2007).

256. Cartegni, L., Chew, S.L. & Krainer, A.R. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* **3**, 285-98 (2002).

257. Inlow, J.K. & Restifo, L.L. Molecular and comparative genetics of mental retardation. *Genetics* **166**, 835-81 (2004).

258. Humeau, Y., Gambino, F., Chelly, J. & Vitale, N. X-linked mental retardation: focus on synaptic function and plasticity. *J Neurochem* **109**, 1-14 (2009).

259. Murre, C., McCaw, P.S., Vaessin, H. *et al.* Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537-44 (1989).

260. de Pontual, L., Mathieu, Y., Golzio, C. *et al.* Mutational, functional, and expression studies of the TCF4 gene in Pitt-Hopkins syndrome. *Hum Mutat* **30**, 669-76 (2009).

261. Giurgea, I., Missirian, C., Cacciagli, P. *et al.* TCF4 deletions in Pitt-Hopkins Syndrome. *Hum.Mutat.* **29**, E242-E251 (2008).

262. Zweier, C., Sticht, H., Bijlsma, E.K. *et al.* Further delineation of Pitt-Hopkins syndrome: phenotypic and genotypic description of 16 novel patients. *J.Med.Genet.* **45**, 738-744 (2008).

263. Abecasis, G.R., Cherny, S.S., Cookson, W.O. & Cardon, L.R. GRR: graphical representation of relationship errors. *Bioinformatics* **17**, 742-3 (2001).

264. Stork, T., Thomas, S., Rodrigues, F. *et al.* Drosophila Neurexin IV stabilizes neuron-glia interactions at the CNS midline by binding to Wrapper. *Development* **136**, 1251-1261 (2009).

265. Zhang, W., Rohlmann, A., Sargsyan, V. *et al.* Extracellular domains of alpha-neurexins participate in regulating synaptic transmission by selectively affecting N- and P/Q-type Ca2+ channels. *J.Neurosci.* **25**, 4330-4342 (2005).

266. Bakkaloglu, B., O'Roak, B.J., Louvi, A. *et al.* Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *Am.J.Hum.Genet.* **82**, 165-173 (2008).

267. Caudy, M., Vassin, H., Brand, M. *et al.* daughterless, a Drosophila gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the achaete-scute complex. *Cell* **55**, 1061-1067 (1988).

268. Wheeler, S.R., Banerjee, S., Blauth, K. *et al.* Neurexin IV and Wrapper interactions mediate Drosophila midline glial migration and axonal ensheathment. *Development* **136**, 1147-1157 (2009).

269. Li, J., Ashley, J., Budnik, V. & Bhat, M.A. Crucial role of Drosophila neurexin in proper active zone apposition to postsynaptic densities, synaptic growth, and synaptic transmission. *Neuron* **55**, 741-755 (2007).

270. Zeng, X., Sun, M., Liu, L. *et al.* Neurexin-1 is required for synapse formation and larvae associative learning in Drosophila. *FEBS Lett.* **581**, 2509-2516 (2007).

271. Koh, Y.H., Gramates, L.S. & Budnik, V. Drosophila larval neuromuscular junction: molecular components and mechanisms underlying synaptic plasticity. *Microsc.Res.Tech.* **49**, 14-25 (2000).

272. Alarcon, M., Abrahams, B.S., Stone, J.L. *et al.* Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *Am.J.Hum.Genet.* **82**, 150-159 (2008).

273. Arking, D.E., Cutler, D.J., Brune, C.W. *et al.* A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am.J.Hum.Genet.* **82**, 160-164 (2008).

274. Bucan, M., Abrahams, B.S., Wang, K. *et al.* Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS.Genet.* **5**, e1000536 (2009).

275. Feng, J., Schroer, R., Yan, J. *et al.* High frequency of neurexin 1beta signal peptide structural variants in patients with autism. *Neurosci.Lett.* **409**, 10-13 (2006).

276. Friedman, J.I., Vrijenhoek, T., Markx, S. *et al.* CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. *Mol.Psychiatry* **13**, 261-266 (2008).

277. Glessner, J.T., Wang, K., Cai, G. *et al.* Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**, 569-573 (2009).

278. Kim, H.G., Kishikawa, S., Higgins, A.W. *et al.* Disruption of neurexin 1 associated with autism spectrum disorder. *Am.J.Hum.Genet.* **82**, 199-207 (2008).

279. Marshall, C.R., Noor, A., Vincent, J.B. *et al.* Structural variation of chromosomes in autism spectrum disorder. *Am.J.Hum.Genet.* **82**, 477-488 (2008).

280. Rujescu, D., Ingason, A., Cichon, S. *et al.* Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum.Mol.Genet.* **18**, 988-996 (2009).

281. Vernes, S.C., Newbury, D.F., Abrahams, B.S. *et al.* A functional genetic link between distinct developmental language disorders. *N.Engl.J.Med.* **359**, 2337-2345 (2008).

282. Yan, J., Noltner, K., Feng, J. *et al.* Neurexin 1alpha structural variants associated with autism. *Neurosci.Lett.* **438**, 368-370 (2008).

283. Verkerk, A.J., Mathews, C.A., Joosse, M. *et al.* CNTNAP2 is disrupted in a family with Gilles de la Tourette syndrome and obsessive compulsive disorder. *Genomics* **82**, 1-9 (2003).

284. Belloso, J.M., Bache, I., Guitart, M. *et al.* Disruption of the CNTNAP2 gene in a t(7;15) translocation family without symptoms of Gilles de la Tourette syndrome. *Eur.J.Hum.Genet.* **15**, 711-713 (2007).

285. Jackman, C., Horn, N.D., Molleston, J.P. & Sokol, D.K. Gene associated with seizures, autism, and hepatomegaly in an Amish girl. *Pediatr Neurol* **40**, 310-3 (2009).

286. Zahir, F.R., Baross, A., Delaney, A.D. *et al.* A patient with vertebral, cognitive and behavioural abnormalities and a de novo deletion of NRXN1alpha. *J.Med.Genet.* **45**, 239-243 (2008).

287. Sudhof, T.C. Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**, 903-911 (2008).

288. Missler, M. & Sudhof, T.C. Neurexins: three genes and 1001 products. *Trends Genet.* **14**, 20-26 (1998).

289. Missler, M., Zhang, W., Rohlmann, A. *et al.* Alpha-neurexins couple Ca2+ channels to synaptic vesicle exocytosis. *Nature* **423**, 939-948 (2003).

290. Laumonnier, F., Bonnet-Brilhault, F., Gomot, M. *et al.* X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am.J.Hum.Genet.* **74**, 552-557 (2004).

291. Durand, C.M., Betancur, C., Boeckers, T.M. *et al.* Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* **39**, 25-7 (2007).

292. Bourgeron, T. A synaptic trek to autism. *Curr Opin Neurobiol* (2009).

293. Bellen, H.J., Lu, Y., Beckstead, R. & Bhat, M.A. Neurexin IV, caspr and paranodin--novel members of the neurexin family: encounters of axons and glia. *Trends Neurosci.* **21**, 444-449 (1998).

294. Arroyo, E.J., Xu, T., Poliak, S. *et al.* Internodal specializations of myelinated axons in the central nervous system. *Cell Tissue Res.* **305**, 53-66 (2001).

295. Poliak, S., Salomon, D., Elhanany, H. *et al.* Juxtaparanodal clustering of Shaker-like K+ channels in myelinated axons depends on Caspr2 and TAG-1. *J.Cell Biol.* **162**, 1149-1160 (2003).

296. Baumgartner, S., Littleton, J.T., Broadie, K. *et al.* A Drosophila neurexin is required for septate junction and blood-nerve barrier formation and function. *Cell* **87**, 1059-1068 (1996).

297. Poliak, S., Gollan, L., Martinez, R. *et al.* Caspr2, a new member of the neurexin superfamily, is localized at the juxtaparanodes of myelinated axons and associates with K+ channels. *Neuron* **24**, 1037-1047 (1999).

298. Wagh, D.A., Rasse, T.M., Asan, E. *et al.* Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in Drosophila. *Neuron* **49**, 833-844 (2006).

299. Ching, M.S., Shen, Y., Tan, W.H. *et al.* Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. *Am J Med Genet B Neuropsychiatr Genet* **153B**, 937-47 (2010).

300. Wisniowiecka-Kowalnik, B., Nesteruk, M., Peters, S.U. *et al.* Intragenic rearrangements in NRXN1 in three families with autism spectrum disorder, developmental delay, and speech delay. *Am J Med Genet B Neuropsychiatr Genet* **153B**, 983-93 (2010).

301. Awadalla, P., Gauthier, J., Myers, R.A. *et al.* Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. *Am J Hum Genet* **87**, 316-24 (2010).

302. Bradley, W.E., Raelson, J.V., Dubois, D.Y. *et al.* Hotspots of large rare deletions in the human genome. *PLoS One* **5**, e9401 (2010).

303. Kirov, G., Gumus, D., Chen, W. *et al.* Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* **17**, 458-65 (2008).

304. Magri, C., Sacchetti, E., Traversa, M. *et al.* New copy number variations in schizophrenia. *PLoS One* **5**, e13422 (2010).

305. Mefford, H.C., Muhle, H., Ostertag, P. *et al.* Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. *PLoS Genet* **6**, e1000962 (2010).

306. Szatmari, P., Paterson, A.D., Zwaigenbaum, L. *et al.* Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* **39**, 319-28 (2007).

307. Vrijenhoek, T., Buizer-Voskamp, J.E., van, d.S.I. *et al.* Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am.J.Hum.Genet.* **83**, 504-510 (2008).

308. Zweier, C., de Jong, E.K., Zweier, M. *et al.* CNTNAP2 and NRXN1 are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in Drosophila. *Am J Hum Genet* **85**, 655-66 (2009).

309. Borozdin, W., Boehm, D., Leipoldt, M. *et al.* SALL4 deletions are a common cause of Okihiro and acro-renal-ocular syndromes and confirm haploinsufficiency as the pathogenic mechanism. *J Med Genet* **41**, e113 (2004).

310. Engels, H., Wohlleber, E., Zink, A. *et al.* A novel microdeletion syndrome involving 5q14.3-q15: clinical and molecular cytogenetic characterization of three patients. *Eur J Hum Genet* **17**, 1592-9 (2009).

311. Evans, D.G., Newton, V., Neary, W. *et al.* Use of MRI and audiological tests in presymptomatic diagnosis of type 2 neurofibromatosis (NF2). *J Med Genet* **37**, 944-7 (2000).

312. Evans, D.G. & Wallace, A. An update on age related mosaic and offspring risk in neurofibromatosis 2 (NF2). *J Med Genet* **46**, 792 (2009).

313. Evans, D.G., Wallace, A.J., Wu, C.L. *et al.* Somatic mosaicism: a common cause of classic disease in tumor-prone syndromes? Lessons from type 2 neurofibromatosis. *Am J Hum Genet* **63**, 727-36 (1998).

314. Flint, J. Genetic basis of cognitive disability. *Dialogues Clin Neurosci* **3**, 37-46 (2001).

315. Forbat, L. An analysis of key principles in Valuing People: implications for supporting people with dementia. *J Intellect Disabil* **10**, 249-60 (2006).

316. Forrest, M., Chapman, R.M., Doyle, A.M. *et al.* Functional analysis of TCF4 missense mutations that cause Pitt-Hopkins syndrome. *Hum Mutat* (2012).

317. Ropers, H.H. Genetics of intellectual disability. *Curr Opin Genet Dev* **18**, 241-50 (2008).

318. van Bokhoven, H. Genetic and Epigenetic Networks in Intellectual Disabilities. *Annual Review of Genetics* **45**, 81-104 (2011).

319. Najmabadi, H., Hu, H., Garshasbi, M. *et al.* Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* **478**, 57-63 (2011).

320. Pavlowsky, A., Chelly, J. & Billuart, P. Emerging major synaptic signaling pathways involved in intellectual disability. *Mol Psychiatry* **17**, 682-93 (2012).

321. Gilman, S.R., Iossifov, I., Levy, D. *et al.* Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron* **70**, 898-907 (2011).

322. Bayes, A., van de Lagemaat, L.N., Collins, M.O. *et al.* Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nat Neurosci* **14**, 19-21 (2011).

323. Dang, V.T., Kassahn, K.S., Marcos, A.E. & Ragan, M.A. Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *Eur J Hum Genet* **16**, 1350-7 (2008).

324. Wu, C., Orozco, C., Boyer, J. *et al.* BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* **10**, R130 (2009).

325. Betancur, C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res* **1380**, 42-77 (2011).

326. Oortveld, M.A., Keerthikumar, S., Oti, M. *et al.* Human intellectual disability genes form conserved functional modules in Drosophila. *PLoS Genet* **9**, e1003911 (2013).

327. Robinson, P.N. & Mundlos, S. The human phenotype ontology. *Clin Genet* **77**, 525-34 (2010).

328. Robinson, P.N., Kohler, S., Bauer, S. *et al.* The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet* **83**, 610-5 (2008).

329. Laumonnier, F., Cuthbert, P.C. & Grant, S.G. The role of neuronal complexes in human X-linked brain diseases. *Am J Hum Genet* **80**, 205-20 (2007).

330. Biesecker, L.G. Lumping and splitting: molecular biology in the genetics clinic. *Clin Genet* **53**, 3-7 (1998).

331. Marangi, G., Ricciardi, S., Orteschi, D. *et al.* The Pitt-Hopkins syndrome: report of 16 new patients and clinical diagnostic criteria. *Am J Med Genet A* **155A**, 1536-45 (2011).

332. Steinbusch, C., van Roozendaal, K., Tserpelis, D. *et al.* Somatic mosaicism in a mother of two children with Pitt-Hopkins syndrome. *Clin Genet* (2012).

333. Whalen, S., Heron, D., Gaillon, T. *et al.* Novel comprehensive diagnostic strategy in Pitt-Hopkins syndrome: clinical score and further delineation of the TCF4 mutational spectrum. *Hum Mutat* **33**, 64-72 (2012).

334. Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F. & Monaco, A.P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519-23 (2001).

335. MacDermot, K.D., Bonora, E., Sykes, N. *et al.* Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet* **76**, 1074-80 (2005).

336. Amiel, J., Laudier, B., Attie-Bitach, T. *et al.* Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. *Nat Genet* **33**, 459-61 (2003).

337. Romeo, G., Ronchetto, P., Luo, Y. *et al.* Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. *Nature* **367**, 377-8 (1994).

338. Edery, P., Lyonnet, S., Mulligan, L.M. *et al.* Mutations of the RET proto-oncogene in Hirschsprung's disease. *Nature* **367**, 378-80 (1994).

339. Sepp, M., Pruunsild, P. & Timmusk, T. Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum Mol Genet* **21**, 2873-88 (2012).

340. Harrison, V., Connell, L., Hayesmoore, J. *et al.* Compound heterozygous deletion of NRXN1 causing severe developmental delay with early onset epilepsy in two sisters. *Am J Med Genet A* **155A**, 2826-31 (2011).

341. Zweier, C. Severe Intellectual Disability Associated with Recessive Defects in CNTNAP2 and NRXN1. *Mol Syndromol* **2**, 181-185 (2012).

342. Marshall, C.R., Noor, A., Vincent, J.B. *et al.* Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* **82**, 477-88 (2008).

343. Steinberg, S., de Jong, S., Andreassen, O.A. *et al.* Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet* **20**, 4076-81 (2011).

344. Duong, L., Klitten, L.L., Moller, R.S. *et al.* Mutations in NRXN1 in a family multiply affected with brain disorders: NRXN1 mutations and brain disorders. *Am J Med Genet B Neuropsychiatr Genet* **159B**, 354-8 (2012).

345. Iqbal, Z., Vandeweyer, G., van der Voet, M. *et al.* Homozygous and heterozygous disruptions of ANK3: at the crossroads of neurodevelopmental and psychiatric disorders. *Hum Mol Genet* **22**, 1960-70 (2013).

346. Gregor, A., Albrecht, B., Bader, I. *et al.* Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1. *BMC Med Genet* **12**, 106 (2011).

347. Schaaf, C.P., Boone, P.M., Sampath, S. *et al.* Phenotypic spectrum and genotype-phenotype correlations of NRXN1 exon deletions. *Eur J Hum Genet* **20**, 1240-7 (2012).

348. Dabell, M.P., Rosenfeld, J.A., Bader, P. *et al.* Investigation of NRXN1 deletions: clinical and molecular characterization. *Am J Med Genet A* **161A**, 717-31 (2013).

349. Bena, F., Bruno, D.L., Eriksson, M. *et al.* Molecular and clinical characterization of 25 individuals with exonic deletions of NRXN1 and comprehensive review of the literature. *Am J Med Genet B Neuropsychiatr Genet* **162B**, 388-403 (2013).

350. Neale, B.M., Kou, Y., Liu, L. *et al.* Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**, 242-5 (2012).

351. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. *Science* **273**, 1516-7 (1996).

352. O'Roak, B.J., Deriziotis, P., Lee, C. *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* **43**, 585-9 (2011).

353. Iossifov, I., Ronemus, M., Levy, D. *et al.* De novo gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285-99 (2012).

354. Girard, S.L., Gauthier, J., Noreau, A. *et al.* Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat Genet* **43**, 860-3 (2011).

355. Xu, B., Roos, J.L., Dexheimer, P. *et al.* Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat Genet* **43**, 864-8 (2011).

356. Hamdan, F.F., Daoud, H., Rochefort, D. *et al.* De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet* **87**, 671-8 (2010).

357. Horn, D., Kapeller, J., Rivera-Brugues, N. *et al.* Identification of FOXP1 deletions in three unrelated patients with mental retardation and significant speech and language deficits. *Hum Mutat* **31**, E1851-60 (2010).

358. Vaags, A.K., Lionel, A.C., Sato, D. *et al.* Rare deletions at the neurexin 3 locus in autism spectrum disorder. *Am J Hum Genet* **90**, 133-41 (2012).

359. Kumar, R.A., KaraMohamed, S., Sudi, J. *et al.* Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* **17**, 628-38 (2008).

360. O'Roak, B.J., Vives, L., Fu, W. *et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* **338**, 1619-22 (2012).

361. Carvill, G.L., Heavin, S.B., Yendle, S.C. *et al.* Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat Genet* **45**, 825-30 (2013).

Acknowledgements

Finishing this thesis is one of my big milestones after a “long” time in human genetics. The study is the result of work at two places. First, the Institute of Human Genetics in Erlangen, where I started 12 years ago with my medical doctoral thesis on Mowat-Wilson syndrome and where I still find excitement in the work I do. In 2009 I spent a fantastic year in the Department of Human Genetics at the Radboud University in Nijmegen, where I learned lots of new things and gained new perspectives.

I am indebted to many people who have accompanied my way so far and made my work and this study possible. The most important are:

My promoters: Prof. André Reis and Prof. Han Brunner. Dear Prof. Reis, thank you for providing a stimulating environment for many years, for your continuous support and for many challenges and opportunities, one of them to spend a year away in Nijmegen. Dear Han, thank you for letting me always feel welcome in your Department and for giving me the great chance to accomplish this thesis.

My copromoter: Dear Annette, thanks for not only a great year in your group but also for your friendship and mentoring since then! Your enthusiasm, knowledge and views have greatly enriched both my work and personal life.

I would also like to thank the members of the manuscript committee: Prof. Hans van Bokhoven, Prof. Michel Willemsen, and Prof. Ype Elgersma. Hans, thank you for letting me be part of your group while I was in Nijmegen. I enjoyed the time a lot! Ype, your suggestion during a long beach walk at the conference in Brazil was the starting point for this thesis. Thanks for this and I hope to meet you at further nice places in the future!

Anita, thank you for introducing me to the world of human genetics and for guiding me during the first years. Your motivation, stamina, and confidence formed me a lot.

The Fly group: Jamie, Eiko, Korinna, Merel, Bonnie, Jolanda, it was a great time with you! Thanks for expanding my household by giving me the two guinea pigs UAS and Gal4 as a good-bye present when I returned to Erlangen. You and also the new fly group members Anna, Misa, Tom and Monique are the main reason my visits to Nijmegen still feel like coming home. Korinna, thanks for always providing me a bed in “Hotel Kochinke”.

My office mates: Jamie, Eiko, Dorus, and Jeroen, I enjoyed sharing the office with you in relaxed and fruitful atmosphere. Also thanks to all other people, especially in the molecular lab on P5 and on P2 like Janneke, Zafar, Margit and Evelyn.

My guesthouse (and office) neighbours: Tony and Michael, who fed me many exquisite dinners! Thank you and also thanks to Thilo for being part of my “dutch” family.

Shiva, Martjin and Martin, thanks for the work on the SysID-database, the list and datasets. Also thanks to Pavel, Misa, Bonnie, Korinna and Jolanda for all the efforts with the database, and to Tjitske for the clinical revisions.

I would also like to acknowledge all the clinicians who contributed to this work by providing samples and clinical information. Working on rare syndromes is only possible with good collaborations and many good clinical eyes.

I would also like to thank my colleagues in Erlangen who have accompanied me for many years and with whom I share lots of fun, fruitful discussions, and regular lunch. Michi, thank you for your artistic contribution to this book. Anne and Christine, you are forming my own group now. Thanks for excellent work and for always being enthusiastic about what you do!

I am grateful to my family and friends for always being supportive of what I decide to do.

Curriculum vitae

Christiane Zweier was born on June 21st, 1978 in Bamberg, Germany. She finished her pre-university education in 1997 at the “Dientzenhofer-Gymnasium” in Bamberg. In the same year she started to study medicine at the Friedrich-Alexander-University in Erlangen. Parallel to her studies she started her medical thesis in the Institute of Human Genetics in the group of Anita Rauch in 2000 with the title: “Mowat-Wilson syndrome: clinical and genetic delineation of a new disease entity”. After finishing university in 2003, she started her clinical training at the Institute of Human Genetics in January 2004 and continued to work in the group of Anita Rauch on the identification of novel genes for intellectual disability and the clinical, molecular and functional characterization of the associated genotypes and phenotypes. In 2009, Christiane spent a year as a guest scientist in Annette Schenck’s group at the Human Genetics Department of the Radboud University Nijmegen and learned to work with *Drosophila melanogaster* as a model organism for intellectual disability. Back in Erlangen, she started her own group and established a *Drosophila* lab on order to further combine clinical, molecular and functional studies on intellectual disability. In 2012 she finished her training in clinical genetics and in 2013 her “habilitation”.

List of publications

52. van Rahden VA, Rau I, Fuchs S, Kosyna FK, de Almeida HL Jr, Fryssira H, Isidor B, Jauch A, Joubert M, Lachmeijer AM, **Zweier C**, Moog U, Kutsche K. **Clinical spectrum of females with HCCS mutation: from no clinical signs to a neonatal lethal form of the microphthalmia with linear skin defects (MLS) syndrome.** Orphanet J Rare Dis. 2014 Apr 15;9(1):53.

51. Vulto-van Silfhout AT, Rajamanickam S, Jensik PJ, Vergult S, de Rocker N, Newhall KJ, Raghavan R, Reardon SN, Jarrett K, McIntyre T, Bulinski J, Ownby SL, Huggenvik JI, McKnight GS, Rose GM, Cai X, Willaert A, **Zweier C**, Endele S, de Ligt J, van Bon BW, Lugtenberg D, de Vries PF, Veltman JA, van Bokhoven H, Brunner HG, Rauch A, de Brouwer AP, Carvill GL, Hoischen A, Mefford HC, Eichler EE, Vissers LE, Menten B, Collard MW, de Vries BB. **Mutations Affecting the SAND Domain of DEAF1 Cause Intellectual Disability with Severe Speech Impairment and Behavioral Problems**. Am J Hum Genet. 2014 May 1;94(5):649-61.

50. Basel-Vanagaite L, Yilmaz R, Tang S, Reuter MS, Rahner N, Grange DK, Mortenson M, Koty P, Feenstra H, Farwell Gonzalez KD, Sticht H, Boddaert N, Désir J, Anyane-Yeboa K, **Zweier C**, Reis A, Kubisch C, Jewett T, Zeng W, Borck G. **Expanding the clinical and mutational spectrum of Kaufman oculocerebrofacial syndrome with biallelic UBE3B mutations**. Hum Genet. 2014 Mar 11.

49. Agha Z, Iqbal Z, Azam M, Siddique M, Willemsen MH, Kleefstra T, **Zweier C**, de Leeuw N, Qamar R, van Bokhoven H. **A complex microcephaly syndrome in a Pakistani family associated with a novel missense mutation in RBBP8 and a heterozygous deletion in NRXN1**. Gene. 2014 Mar 15;538(1):30-5.

48. Oortveld MA, Keerthikumar S, Oti M, Nijhof B, Fernandes AC, Kochinke K, Castells-Nobau A, van Engelen E, Ellenkamp T, Eshuis L, Galy A, van Bokhoven H, Habermann B, Brunner HG, **Zweier C**, Verstreken P, Huynen MA, Schenck A. **Human intellectual disability genes form conserved functional modules in Drosophila.** PLoS Genet. 2013 Oct;9(10):e1003911.

47. Hofmann K, Zweier M, Sticht H, **Zweier C**, Wittmann W, Hoyer J, Uebe S, van Haeringen A, Thiel CT, Ekici AB, Reis A, Rauch A. **Biallelic SEMA3A defects cause a novel type of syndromic short stature.** Am J Med Genet A. 2013 Nov;161(11):2880-9.

46. **Zweier C**, Kraus C, Brueton L, Cole T, Degenhardt F, Engels H, Gillessen-Kaesbach G, Graul-Neumann L, Horn D, Hoyer J, Just W, Rauch A, Reis A, Wollnik B, Zeschnigk M, Lüdecke HJ, Wieczorek D. **A new face of Borjeson-Forssman-Lehmann syndrome? De novo mutations in PHF6 in seven females with a distinct phenotype.** J Med Genet. 2013 Dec;50(12):838-47.

45. Wieczorek D, Bögershausen N, Beleggia F, Steiner-Haldenstätt S, Pohl E, Li Y, Milz E, Martin M, Thiele H, Altmüller J, Alanay Y, Kayserili H, Klein-Hitpass L, Böhringer S, Wollstein A, Albrecht B, Boduroglu K, Caliebe A, Chrzanowska K, Cogulu O, Cristofoli F, Czeschik JC, Devriendt K, Dotti MT, Elcioglu N, Gener B, Goecke TO, Krajewska-Walasek M, Guillén-Navarro E, Hayek J, Houge G, Kilic E, Simsek-Kiper PO, López-González V, Kuechler A, Lyonnet S, Mari F, Marozza A, Mathieu Dramard M, Mikat B, Morin G, Morice-Picard F, Ozkynay F, Rauch A, Renieri A, Tinschert S, Utine GE, Vilain C, Vivarelli R, **Zweier C**, Nürnberg P, Rahmann S, Vermeesch J, Lüdecke HJ, Zeschnigk M, Wollnik B. **A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling.** Hum Mol Genet. 2013 Aug 14. [Epub ahead of print]

44. Gregor A, Oti M, Kouwenhoven EN, Hoyer J, Sticht H, Ekici AB, Kjaergaard S, Rauch A, Stunnenberg HG, Uebe S, Vasileiou G, Reis A, Zhou H, **Zweier C**. **De Novo Mutations in the Genome Organizer CTCF Cause Intellectual Disability.** Am J Hum Genet. 2013 Jul

11;93(1):124-31

43. Makrythanasis P, van Bon BW, Steehouwer M, Rodríguez-Santiago B, Simpson M, Dias P, Anderlid BM, Arts P, Bhat M, Augello B, Biamino E, Bongers EM, Del Campo M, Cordeiro I, Cueto-González AM, Cuscó I, Deshpande C, Frysira E, Louise Izatt, Flores R, Galán E, Gener B, Gilissen C, Granneman SM, Hoyer J, Yntema HG, Kets CM, Koolen DA, Marcelis CL, Medeira A, Micale L, Mohammed S, de Munnik SA, Nordgren A, Reardon SP, Revencu N, Roscioli T, Ruiterkamp-Versteeg M, Santos HG, Schoumans J, Schuurs-Hoeijmakers JH, Silengo MC, Toledo L, Vendrell T, van der Burgt I, van Lier B, **Zweier C**, Reymond A, Trembath RC, Perez-Jurado L, Dupont J, de Vries BB, Brunner HG, Veltman JA, Merla G, Antonarakis SE, Hoischen A. **MLL2 mutation detection in 86 patients with Kabuki syndrome: a genotype-phenotype study.** Clin Genet. 2013 Jan 16. doi: 10.1111/cge.12081. [Epub ahead of print]

42. Steinbusch CV, van Roozendaal KE, Tserpelis D, Smeets EE, Kranenburg-de Koning TJ, de Waal KH, **Zweier C**, Rauch A, Hennekam RC, Blok MJ, Schrander-Stumpel CT. **Somatic mosaicism in a mother of two children with Pitt-Hopkins syndrome.** Clin Genet. 2013 Jan;83(1):73-77

41. Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, Albrecht B, Bartholdi D, Beygo J, Di Donato N, Dufke A, Cremer K, Hempel M, Horn D, Hoyer J, Joset P, Röpke A, Moog U, Riess A, Thiel CT, Tzschach A, Wiesener A, Wohlleber E, **Zweier C**, Ekici AB, Zink AM, Rump A, Meisinger C, Grallert H, Sticht H, Schenck A, Engels H, Rappold G, Schröck E, Wieacker P, Riess O, Meitinger T, Reis A, Strom TM. **Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study.** Lancet. 2012 Nov 10;380(9854):1674-82.

40. Gregor A, Krumbiegel M, Kraus C, Reis A, **Zweier C**. **De Novo Triplication of the MAPT Gene from the Recurrent 17q21.31 Microdeletion Region in a Patient with Moderate Intellectual Disability and Various Minor Anomalies.** Am J Med Genet A. 2012 Jul;158A(7):1765-70.

39. Wolff D, Endele S, Azzarello-Burri S, Hoyer J, Zweier M, Schanze I, Schmitt B, Rauch A, Reis A, **Zweier C**. **In frame deletion and missense mutations of the C-terminal helicase domain of SMARCA2 in three patients with Nicolaides-Baraitser syndrome.** Mol Syndromol. 2012 Apr;2(6):237-244.

38. Armani R, Archer H, Clarke A, Vasudevan P, **Zweier C**, Ho G, Williamson S, Cloosterman D, Yang N, Christodoulou J. **Transcription Factor 4 and Myocyte Enhancer Factor 2C mutations are not common causes of Rett syndrome.** Am J Med Genet A. 2012 Apr;158A(4):713-9.

37. **Zweier C**. **Severe Intellectual Disability Associated with Recessive Defects in CNTNAP2 and NRXN1.** Mol Syndromol. 2012 Apr;2(3-5):181-185.

36. Hoyer J, Ekici AB, Endele S, Popp B, **Zweier C**, Wiesener A, Wohlleber E, Dufke A, Petsch C, Zweier M, Göhring I, Zink AM, Rappold G, Schröck E, Wieczorek D, Riess O, Engels H, Rauch A, Reis A. **Haploinsufficiency of ARID1B, a member of the SWI/SNF-A chromatin-remodeling complex, is a frequent cause of intellectual disability.** Am J Hum Genet. 2012 Mar 9;90(3):565-572.

35. Gregor A, Albrecht B, Bader I, Bijlsma EK, Ekici AB, Engels H, Hackmann K, Horn D, Hoyer J, Klapecki J, Kohlhase J, Maystadt I, Nagl S, Prott E, Tinschert S, Ullmann R, Wohlleber E, Woods G, Reis A, Rauch A, **Zweier C**. **Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1.** BMC Med Genet. 2011 Aug 9;12:106.

34. Busche A, Graul-Neumann LM, **Zweier C**, Rauch A, Klopocki E, Horn D. **Microdeletions of chromosome 7p21, including TWIST1, associated with significant microcephaly, facial**

dysmorphism, and short stature. Eur J Med Genet. 2011 May-Jun;54(3):256-61. Epub 2011 Feb 17.

33. Dauwerse JG, Dixon J, Seland S, Ruivenkamp CA, van Haeringen A, Hoefsloot LH, Peters DJ, Boers AC, Daumer-Haas C, Maiwald R, **Zweier C**, Kerr B, Cobo AM, Toral JF, Hoogeboom AJ, Lohmann DR, Hehr U, Dixon MJ, Breuning MH, Wieczorek D. **Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome.** Nat Genet. 2011 Jan;43(1):20-2.

32. Griffin HR, Töpf A, Glen E, **Zweier C**, Stuart AG, Parsons J, Peart I, Deanfield J, O'Sullivan J, Rauch A, Scambler P, Burn J, Cordell HJ, Keavney B, Goodship JA. **Systematic survey of variants in TBX1 in non-syndromic tetralogy of Fallot identifies a novel 57 base pair deletion that reduces transcriptional activity but finds no evidence for association with common variants.** Heart. 2010 Oct;96(20):1651-5.

31. Zweier M, Gregor A, **Zweier C**, Engels H, Sticht H, Wohlleber E, Bijlsma EK, Holder SE, Zenker M, Rossier E, Grasshoff U, Johnson DS, Robertson L, Firth HV; Cornelia Kraus, Ekici AB, Reis A, Rauch A. **Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression.** Hum Mutat. 2010 Jun;31(6):722-33.

30. Smigiel R, Szafranska A, Czyzewska M, Rauch A, **Zweier Ch**, Patkowski D. **Severe clinical course of Hirschsprung disease in a Mowat-Wilson syndrome patient.** J Appl Genet. 2010;51(1):111-3.

29. Rauch R, Hofbeck M, **Zweier C**, Koch A, Zink S, Hoyer J, Kaulitz R, Singer H, Rauch A. Comprehensive genotype-phenotype analysis in 230 patients with tetralogy of Fallot. J Med Genet. 2010 May;47(5):321-31.

28. Vandewalle J, Van Esch H, Govaerts K, Verbeeck J, **Zweier C**, Madrigal I, Mila M, Pijkels E, Fernandez I, Kohlhase J, Spaich C, Rauch A, Fryns JP, Marynen P, Froyen G. **Dosage-Dependent Severity of the Phenotype in Patients with Mental Retardation Due to a Recurrent Copy-Number Gain at Xq28 Mediated by an Unusual Recombination.** Am J Hum Genet. 2009 Dec; 85(6):809-822

27. **Zweier C**, de Jong EK, Zweier M, Orrico A, Ousager LB, Collins AL, Bijlsma EK, Oortveld MA, Ekici AB, Reis A, Schenck A, Rauch A. **CNTNAP2 and NRXN1 are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in Drosophila.** Am J Hum Genet. 2009 Nov;85(5):655-66.

26. Heinritz W, Hüffmeier U, Strenge S, Miterski B, **Zweier C**, Leinung S, Bohring A, Mitulla B, Peters U, Froster UG. New mutations of EXT1 and EXT2 genes in German patients with Multiple Osteochondromas. Ann Hum Genet. 2009 May;73(Pt 3):283-91.

25. Cisse B, Caton ML, Lehner M, Maeda T, Scheu S, Locksley R, Holmberg D, **Zweier C**, den Hollander NS, Kant SG, Holter W, Rauch A, Zhuang Y, Reizis B. **Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development.** Cell, 2008 Oct 3;135(1):37-48.

24. **Zweier C**, Sticht H, Bijlsma E, Clayton-Smith J, Boonen SE, Fryer A, Greally MT, Hoffmann L, den Hollander NS, Jongmans M, Kant SG, King MD, Lynch SA, McKee S, Midro AT, Park SM, Ricotti V, Tarantino E, Wessels M, Peippo M, Rauch A. **Further delineation of Pitt-Hopkins syndrome: phenotypic and genotypic description of sixteen novel patients.** J Med Genet. 2008 Nov;45(11):738-44

23. **Zweier C**, Trautmann U, Ekici A, Rauch A. **A 15Mb duplication of 6q24.1-q25.3 associated with typical but milder features of the duplication 6q syndrome.** Eur J Med Genet. 2008 Jul-Aug;51(4):358-61.

22. Rauch A, Thiel CT, Schindler D, Wick U, Crow YJ, Ekici AB, van Essen AJ, Goecke TO, Al-Gazali L, Chrzanowska KH, **Zweier C**, Brunner HG, Becker K, Curry CJ, Dallapiccola B, Devriendt K, Dörfler A, Kinning E, Megarbane A, Meinecke P, Semple RK, Spranger S, Toutain A, Trembath RC, Voß E, Wilson L, Hennekam R, de Zegher F, Dörr HG, Reis A. **Mutations in the Pericentrin (PCNT) Gene Cause Primordial Dwarfism.** Science. 2008 Feb 8;319(5864):816-9.

21. Huehne K, **Zweier C**, Raab K, Odent S, Bonnaure-Mallet M, Sixou JL, Landrieu P, Goizet C, Sarlangue J, Baumann M, Eggermann T, Rauch A, Ruppert S, Stettner GM, Rautenstrauss B. **Novel missense, insertion and deletion mutations in the neurotrophic tyrosine kinase receptor type 1 gene (NTRK1) associated with congenital insensitivity to pain with anhidrosis.** Neuromuscul Disord. 2008 Feb;18(2):159-66.

20. Hoyer J, Dreweke A, Becker C, Göhring I, Thiel CT, Peippo MM, Rauch R, Hofbeck M, Trautmann U, **Zweier C**, Zenker M, Hüffmeier U, Kraus C, Ekici AB, Rüschendorf F, Nürnberg P, Reis A, Rauch. **Molecular karyotyping in patients with mental retardation using 100K single-nucleotide polymorphism arrays.** J Med Genet. 2007 Oct;44(10):629-36.

19. Zenker M, Horn D, Wieczorek D, Allanson J, Pauli S, van der Burgt I, Doerr HG, Gaspar H, Hofbeck M, Gillessen-Kaesbach G, Koch A, Meinecke P, Mundlos S, Nowka A, Rauch A, Reif S, von Schnakenburg C, Seidel H, Wehner LE, **Zweier C**, Bauhuber S, Matejas V, Kratz CP, Thomas C, Kutsche K. **SOS1 is the second most common Noonan gene but plays no major role in cardio-facio-cutaneous syndrome.** J Med Genet. 2007 Oct;44(10):651-6.

18. Strenge S, Heinritz W, **Zweier C**, Rauch A, Rolle U, Merkenschlager A, Froster UG. **Pulmonary artery sling and congenital tracheal stenosis in another patient with Mowat-Wilson syndrome.** Am J Med Genet A. 2007 Jul 1;143(13):1528-30.

17. **Zweier C**, Peippo MM, Hoyer J, Sousa S, Bottani A, Clayton-Smith J, Reardon W, Saraiva J, Cabral A, Gohring I, Devriendt K, de Ravel T, Bijlsma EK, Hennekam RC, Orrico A, Cohen M, Dreweke A, Reis A, Nurnberg P, Rauch A. **Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome).** Am J Hum Genet. 2007 May;80(5):994-1001.

16. **Zweier C**, Sticht H, Aydin-Yaylagül I, Campbell CE, Rauch A. **Human TBX1 missense mutations cause gain of function resulting in the same phenotype as 22q11.2 deletions.** Am J Hum Genet. 2007 Mar;80(3):510-7.

15. Rauch A, Hoyer J, Guth S, **Zweier C**, Kraus C, Becker C, Zenker M, Huffmeier U, Thiel C, Ruschendorf F, Nurnberg P, Reis A, Trautmann U. **Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation.** Am J Med Genet A. 2006 Oct 1;140(19):2063-74.

14. Heinritz W, **Zweier C**, Froster UG, Strenge S, Kujat A, Syrbe S, Rauch A, Schuster V. **A missense mutation in the ZFHX1B gene associated with an atypical Mowat-Wilson syndrome phenotype.** Am J Med Genet A. 2006 Jun 1;140(11):1223-7.

13. Jung R, Rauch A, Salomons GS, Verhoeven NM, Jakobs C, Gibson KM, Lachmann E, Sass JO, Trautmann U, **Zweier C**, Staatz G, Knerr I. **Clinical, cytogenetic and molecular characterization of a patient with combined succinic semialdehyde dehydrogenase deficiency and incomplete WAGR syndrome with obesity.** Mol Genet Metab. 2006 Jul;88(3):256-60.

12. **Zweier C**, Horn D, Kraus C, Rauch A. **Atypical ZFHX1B mutation associated with a mild Mowat-Wilson syndrome phenotype.** Am J Med Genet A. 2006 Apr 15;140(8):869-72.

11. Hoornaert KP, Dewinter C, Vereecke I, Beemer FA, Courtens W, Fryer A, Fryssira H, Lees M, Mullner-Eidenbock A, Rimoin D, Siderius L, Superti-Furga A, Temple K, Willems P, Zankl A,

Zweier C, De Paepe A, Coucke P, Mortier GR. **The phenotypic spectrum in patients with arginine to cysteine mutations in the COL2A1 gene.** J Med Genet. 2006 May;43(5):406-13.

10. Thiel CT, Horn D, Zabel B, Ekici AB, Salinas K, Gebhart E, Ruschendorf F, Sticht H, Spranger J, Muller D, **Zweier C**, Schmitt ME, Reis A, Rauch A. **Severely incapacitating mutations in patients with extreme short stature identify RNA-processing endoribonuclease RMRP as an essential cell growth regulator.** Am J Hum Genet. 2005 Nov;77(5):795-806.

09. **Zweier C**, Guth S, Schulte-Mattler U, Rauch A, Trautmann U. **9 Mb deletion including chromosome band 3q24 associated with unsuspicious facial gestalt, persistent ductus omphaloentericus, mild mental retardation and tic.** Eur J Med Genet. 2005 Jul-Sep;48(3):360-2.

08. **Zweier C**, Thiel CT, Dufke A, Crow YJ, Meinecke P, Suri M, Ala-Mello S, Beemer F, Bernasconi S, Bianchi P, Bier A, Devriendt K, Dimitrov B, Firth H, Gallagher RC, Garavelli L, Gillessen-Kaesbach G, Hudgins L, Kaariainen H, Karstens S, Krantz I, Mannhardt A, Medne L, Mucke J, Kibaek M, Krogh LN, Peippo M, Rittinger O, Schulz S, Schelley SL, Temple IK, Dennis NR, Van der Knaap MS, Wheeler P, Yerushalmi B, Zenker M, Seidel H, Lachmeijer A, Prescott T, Kraus C, Lowry RB, Rauch A. **Clinical and mutational spectrum of Mowat-Wilson syndrome.** Eur J Med Genet. 2005 Apr-Jun;48(2):97-111.

07. Garavelli L, Cerruti-Mainardi P, Virdis R, Pedori S, Pastore G, Godi M, Provera S, Rauch A, **Zweier C**, Zollino M, Banchini G, Longo N, Mowat D, Neri G, Bernasconi S. **Genitourinary anomalies in Mowat-Wilson syndrome with deletion/mutation in the zinc finger homeo box 1B gene (ZFHX1B). Report of three Italian cases with hypospadias and review.** Horm Res. 2005;63(4):187-92.

06. Rauch A, Zink S, **Zweier C**, Thiel CT, Koch A, Rauch R, Lascorz J, Huffmeier U, Weyand M, Singer H, Hofbeck M. **Systematic assessment of atypical deletions reveals genotype-phenotype correlation in 22q11.2.** J Med Genet. 2005 Nov;42(11):871-6.

05. Cerruti Mainardi P, Pastore G, **Zweier C**, Rauch A. **Mowat-Wilson syndrome and mutation in the zinc finger homeo box 1B gene: a well defined clinical entity.** J Med Genet. 2004 Feb;41(2):e16.

04. Horn D, Weschke B, **Zweier C**, Rauch A. **Facial phenotype allows diagnosis of Mowat-Wilson syndrome in the absence of Hirschsprung disease.** Am J Med Genet A. 2004 Jan 1;124(1):102-4.

03. **Zweier C**, Temple IK, Beemer F, Zackai E, Sagie T, Weschke B, Rauch A (2003): **Characterization of deletions of the ZFHX1B region and genotype-phenotype analysis in Mowat-Wilson syndrome.** J Med Genet. 2003 Aug;40(8):601-5.

02. Garavelli L, Donadio A, Zanacca C, Banchini G, Della Giustina E, Bertani G, Albertini G, Del Rossi C, Rauch A, **Zweier C**, Zollino M, Neri G. **Hirschsprung disease, mental retardation, characteristic facial features, and mutation in the gene ZFHX1B (SIP1): confirmation of the Mowat-Wilson syndrome.** Am J Med Genet A. 2003 Feb 1;116(4):385-8.

01. **Zweier C**, Albrecht B, Mitulla B, Behrens R, Beese M, Gillessen-Kaesbach G, Rott HD, Rauch A. **"Mowat-Wilson" Syndrome with and without Hirschsprung Disease is a distinct, recognizable Multiple Congenital Anomalies-Mental Retardation Syndrome caused by Mutations in the Zinc finger homeobox 1 B gene (ZFHX1B).** Am J Med Genet. 2002 Mar 15;108(3):177-81.

List of abbreviations

AD	autosomal dominant
AR	autosomal recessive
ASD	autism spectrum disorder
ATP	Adenosine-5'-triphosphate
BAC	bacterial artificial chromosome
Bp	base pair
ca.	circa
ChIP	chromatin immunoprecipitation
arrayCGH	microarray-based comparative genomic hybridization
CM	classic mild to severe or variable
CNV	copy number variant
Ct	threshold cycles
CS	classic severe
dbSNP	database of single nucleotide polymorphisms
DECIPHER	database of chromosomal imbalance and phenotype in humans using ensemble resources
DNA	deoxyribonucleic acid
EBV	Epstein-Barr virus
EEG	electroencephalogram
e.g.	for example
etc.	et cetera
FISH	fluorescence in situ hybridisation
GNF	Genomics Institute of the Novartis Research Foundation
HI	haploinsufficiency
hPSD	post-synaptic density from human neocortex
i.e.	that is
iPS cells	induced pluripotent stem cells
ID	intellectual disability
IQ	Intelligence quotient
kb	kilobase (thousand base pairs)
LD	linkage disequilibrium
LOD	logarithm of odds
Mb	megabase (million base pairs)
MLPA	multiplex ligation dependent probe amplification
MR	mental retardation
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
NC	non-classic
NCBI	national center for biotechnology information
NGS	next generation sequencing
NMD	nonsense-mediated mRNA decay
NMJ	neuromuscular junction
NS	non-syndromic
OFC	occipitofrontal head circumference
OMIM	online mendelian inheritance in man database
ORF	open reading frame
PCR	polymerase chain reaction
qPCR	quantitative polymerase chain reaction
RNA	ribonucleic acid
RNAi	inducible RNA interference

RT-PCR	real time polymerase chain reaction or reverse transcriptase polymerase chain reaction
SD	standard deviation
shRNA	short hairpin RNA
siRNA	short interference RNA
SNP	single nucleotide polymorphism
SWSM	syndromic with structural malformations
SWOSM	syndromic without structural malformations
UAS	upstream activating sequence
UCSC	University of California, Santa Cruz
UTR	untranslated region
WT	wild type